DOCUMENT RESUME

ED 480 977 CE 085 540

AUTHOR	Curtis, David; Denton, Rob
TITLE	The Authentic Performance-Based Assessment of Problem Solving.
INSTITUTION	National Centre for Vocational Education Research, Leabrook (Australia).
ISBN	ISBN-1-74096-167-6
PUB DATE	2003-08-00
NOTE	96p.; Project was a collaborative venture between the Centre for Lifelong Learning and Development and the Torrens Valley Institute of Technical and Further Education (TAFE), Adelaide, Australia.
AVAILABLE FROM	National Centre for Vocational Education Research, 252 Kensington Road, Leabrook, South Australia 5068, Australia (Cat. no. 1316; Web only). Tel: 08 8333 8400; Fax: 08 8331 9211; e-mail: vet_req@ncver.edu.au; Web site: http://www.ncver.edu.au. For full text: http://www.ncver.edu.au/ research/proj/nr2007.pdf.
PUB TYPE	Reports - Research (143) -- Tests/Questionnaires (160)
EDRS PRICE	EDRS Price MF01 Plus Postage. PC Not Available from EDRS
DESCRIPTORS	Check Lists; *Competency Based Education; Critical Thinking; Definitions; Educational Policy; *Evaluation Methods; Guidelines; Information Technology; Literature Reviews; *Performance Based Assessment; Policy Formation; Postsecondary Education; *Problem Solving; Questionnaires; Records (Forms); Secondary Education; Student Attitudes; *Student Evaluation; *Test Construction; Thinking Skills
IDENTIFIERS	*Australia; TAFE (Australia)

ABSTRACT

A new authentic performance-based approach to assessing problem solving was developed for use in vocational education and other programs in Australia. The process of developing the problem-solving assessment instrument and process included the following phases: (1) exploration of the theoretical conceptions of problem solving; (2) identification of five major component processes in problem solving; (3) formulation of sets of indicators for each of the five processes; and (4) description of sets of performance levels for each indicator. A study of the problem-solving assessment instrument established its reliability in assessing problem-solving performance across a range of tasks within the Electronics and Information Technology program at Torrens Valley Institute of Technical and Further Education. The instrument worked well in a validation study undertaken in Torrens Valley Institute's Certificate IV in Workplace Assessment and Training course. The new assessment processes that were used in the trial of the problem-solving assessment instrument, which involved both self-assessment and lecturer validation, not only led to assessment of the key competency of problem solving but also to its development among participants. (Twenty-seven tables/figures are included. The bibliography lists 86 references. The following items are appended: consent and personal details forms; an outline of the problem-solving assessment procedure; the problem-solving inventory; the problem-solving assessment instrument; and student evaluation comments.) (MN)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 480 977

NCVER

The authentic performance-based assessment of problem-solving

genericskills

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY
S. Hayden
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
2A

David Curtis

Rob Denton

BEST COPY AVAILABLE

© **Australian National Training Authority, 2003**

This work has been produced by the National Centre for Vocational Education Research (NCVER) with the assistance of funding provided by the Australian National Training Authority (ANTA). It is published by NCVER under licence from ANTA. Apart from any use permitted under the Copyright Act 1968, no part of this publication may be reported by any process without the written permission of NCVER Ltd. Requests should be made in writing to NCVER Ltd.

The views and opinions expressed in this document are those of the author/project team and do not necessarily reflect the views of the Australian National Training Authority.

ISBN 1 74096 167 6 web edition
TD/TNC 74.03

Published by NCVER
ABN 87 007 967 311
Level 11, 33 King William Road
PO Box 8288, Station Arcade, SA 5000, Australia
http://www.ncver.edu.au

Contents

Tables and figures 5

Acknowledgements 6

Reference group membership 7

Executive summary 8

1 Introduction 11
- Background to the study 11
- The context 12
- Terminology 14

2 Literature review 15
- Key competencies in Australia 15
- The assessment of key competencies 22
- Problem-solving 31
- The assessment of problem-solving 35
- Research questions 36

3 Methods 37
- Project development and approvals 37
- Recruitment and characteristics of participants 38
- Selection and development of instruments 38
- Data collection and analysis 43

4 Results 46
- Participants 46
- The problem-solving inventory 47
- The problem-solving assessment 48
- The validation study 57
- The development of problem-solving ability 58
- Problem-solving and educational achievement 60
- Evaluation by students 62

5 Discussion of results 66
- Participants 66
- The problem-solving inventory 66
- The problem-solving assessment tool 67
- The development of problem-solving ability 68
- Problem-solving and educational achievement 68
- Evaluation by students 69

6 Conclusions and future directions 70
- The feasibility of employing a general-purpose problem-solving assessment 70
- The development of problem-solving ability 72
- The relationships between attitude, problem-solving ability and educational achievement 72
- Future directions 72

BEST COPY AVAILABLE

References 77

Appendices 81

1 Consent and personal details forms 82

2 The problem-solving assessment procedure 84

3 The problem-solving inventory (modified) 86

4 The problem-solving assessment instrument 88

5 Student evaluation comments 92

Tables and figures

Tables

1 Summary of the sets of generic skills proposed by the Australian Industry Group 20
2 Summary of the Australian Chamber of Commerce and Industry/ Business Council of Australia employability skills framework 22
3 Performance levels of indicators using the SOLO taxonomy 41
4 Major processes, indicators and abbreviations of the problem-solving assessment instrument 48
5 Rotated factor solution for the problem-solving assessment 49
6 Results of scale reliabilities 50
7 Scale reliability indices for the PSA and subscales 50
8 Estimates of PSA item locations and thresholds 52
9 Student problem-solving ability (logits and transformed values) 56
10 Estimates of item locations for all data, E&TT data and validation data 58
11 Correlations between attitude, problem-solving ability and educational achievement 61
12 Perceptions of the adequacy of information about key competencies assessments (participants) 63
13 Reasons for seeking key competencies assessments (participants) 63
14 Other perceptions of key competencies assessments (participants) 63
15 Perceptions of the key competencies assessment process using the problem-solving assessment instrument 64
16 Perceptions of the adequacy of information about key competencies assessments (non-participants) 64
17 Classification of comments made by students 65

Figures

1 Fit parameters (Infit MS) of the problem-solving assessment scale 51
2 Category probability curve for item 1 53
3 Category probability curve for item 12 53
4 Distributions of person ability (top) and item thresholds for the PSA 54
5 Participants' problem-solving ability for all assessments on the PS500 scale 55
6 Fit parameters (Infit MS) of the problem-solving assessment in the validation study 57
7 Change in problem-solving performance over two assessment occasions 59
8 Change in problem-solving performance over three assessment occasions 59
9 Path model showing the relationships between attitude, problem-solving ability and educational achievement 62
10 Simulated key competencies profile for inclusion in a report 76

Acknowledgements

The research project reported here is a collaborative venture between the Centre for Lifelong Learning and Development and the Torrens Valley Institute of Technical and Further Education (TAFE), Adelaide, Australia. The researchers are grateful for the support received from both organisations in the conduct of the project. In particular, the support of the Director of Torrens Valley Institute, Virginia Battye, and of the Assistant Director, Nancye Stanelis, is acknowledged. The support and encouragement of the Executive Director of the Centre for Lifelong Learning and Development, Professor Denis Ralph, is also acknowledged.

Many students enrolled in programs at Torrens Valley Institute of TAFE willingly participated in the study and provided helpful feedback to the researchers. Without that participation, the research would not have been possible.

Staff in the Electronics and Information Technology teaching team and in Business Services at Torrens Valley Institute also contributed enthusiastically to this research project and committed additional time from busy teaching schedules to be involved and to support student participation.

Members of the Torrens Valley Institute of TAFE Key Competencies Focus Group took an active interest in the project and encouraged the research activity. Its members have indicated a desire to be involved in other key competencies assessment projects that will be undertaken following this phase of the research.

David Curtis is particularly indebted to Professor Michael Lawson for stimulating an interest in problem-solving and to Professor John Keeves for promoting an interest in current approaches to measurement. The confluence of these two lines of research is reflected in this project.

A reference group was established to monitor the progress of this research project. They provided useful advice and encouragement. They read and commented on interim reports and read several drafts of the final report.

Reference group membership

Gail Jackman	Student representative and member of Student Representative Council, Torrens Valley Institute of TAFE
Professor Michael Lawson	Associate Dean: Research, School of Education, Flinders University
Trevor Parrott	Lecturer, Electronics and Information Technology, Torrens Valley Institute of TAFE
Peter Petkovic	Lecturer, Business Services, Torrens Valley Institute of TAFE
Professor Denis Ralph	Executive Director, Centre for Lifelong Learning and Development
Greg Stillwell	Lecturer, Electronics and Information Technology, Torrens Valley Institute of TAFE
David D Curtis	Research Associate, Centre for Lifelong Learning and Development (researcher)
Rob Denton	Advanced Skills Lecturer, Electronics and Information Technology, Torrens Valley Institute of TAFE (researcher)

Executive summary

A new approach to assessing problem-solving

In this project, titled the Authentic Performance-based Assessment of Problem-Solving Project, a new approach to the assessment of problem-solving has been developed. This approach is embodied in an assessment tool, the problem-solving assessment instrument, and in the method of its administration. The instrument was trialled in the study reported in this document.

The approach is authentic, as it is undertaken on participants' attempts to solve problems that occur routinely in their courses and on tasks that simulate tasks that are expected to be encountered in the workplace.

The assessment is performance based. In attempting routine tasks, learners are evaluated on the extent to which they use and can demonstrate identified problem-solving processes.

The approach is also evidence based. Participants are required to present evidence to show that they have employed identified problem-solving processes, and the assessment made on the quality of the evidence that participants can show.

The assessment is criterion based, since performance criteria are specified in advance for each of the component processes that are identified. The task of both the learner and the assessor is to interpret the evidence against specified performance levels for each process indicator that is described in the problem-solving assessment instrument.

Development of the problem-solving assessment instrument

The problem-solving assessment instrument was developed on the basis of several convergent theories of problem-solving. Its development followed four distinct phases:

- Theoretical conceptions of problem-solving were explored in order to arrive at a coherent description of problem-solving. This provided a basis for the claim of construct validity for the concept as it was implemented.
- Five major component processes in problem-solving were identified from a variety of theoretical positions. These major processes served to delineate the scope of problem-solving and to establish a basis for the content validity of this implementation.
- For each of the five major processes identified, a set of indicators was proposed. These indicators operationalised the major processes and linked the theoretical foundation of the problem-solving assessment instrument to its practical implementation.
- Finally, for each indicator, a set of performance levels was described. These levels provided a basis for scoring the evidence that learners presented to support their claims of the use of problem-solving processes.

Scope and structure of the report

The purpose of this study was to demonstrate that developments in cognitive theories of problem-solving, and in assessment, measurement and reporting, can form the basis of valid assessments of problem-solving performance. A further purpose was to investigate the relationship between demonstrated problem-solving performance and learning within a course.

The report begins with an introduction to the context in which the trial of the problem-solving assessment instrument was conducted (chapter 1). The project was carried out in the Electronics and Information Technology Program at Torrens Valley Institute of TAFE, Adelaide.

The literature review (chapter 2) of this report is extensive. It begins with a review of the emergence of key competencies in Australia. The processes that led to the definition of key competencies are an important factor in understanding some of the issues that remain problematic in this area of public policy. Key issues that have emerged from this analysis are the definition of key competencies and, in particular, their assessment, reporting and certification. Four main approaches that have been taken to the assessment of problem-solving are reviewed. Each is shown to have advantages and to meet particular needs. A broad policy approach that endorses the complementary use of several of these approaches is suggested. Several unresolved issues in their evaluation include the validity, reliability and precision, and forms of reporting, of the assessment of key competencies. This project is an attempt to suggest possible solutions to some of these issues, specifically in the area of problem-solving.

The methods employed in this research (chapter 3) have been strongly influenced by relatively new approaches to measurement. Past practice has deemed as satisfactory the assignment of numerical grades to levels of performance. However, current methods, most notably item response theory, permit the measurement properties of assessment instruments to be tested and linear measurement scales of known precision to be established. Data collected from the administration of the problem-solving assessment instrument were analysed using the Rasch measurement model. These analyses revealed that the instrument does yield satisfactory measures of problem-solving ability. Knowing the precision and the distribution of learners' problem-solving abilities enabled a number of discriminable levels of performance to be established.

The assessment method employed in this project involved two stages: self-assessment by students using the problem-solving assessment instrument, followed by validation by lecturers, also using the instrument. This approach is suggested as a means both of assessing learner performance on problem-solving and also of enhancing this generic ability and learners' explicit knowledge about it.

In addition to the primarily quantitative data that were collected, qualitative data arising from an evaluation by students of the problem-solving assessment instrument and its method of administration were also gathered and analysed. These suggest that the instrument achieved an important purpose—that the assessment process was also a significant learning activity for students.

The analysis of the data was necessarily technical, and its details are presented in chapter 4, Results. However, the results have clear implications for policy makers and practitioners in the field. These matters are discussed separately, in chapter 5, Discussion of results.

In chapter 6, Conclusions and future directions, a consolidated summary of the results of this project is presented. The instrument trialled in this project, the problem-solving assessment, shows considerable promise as a general-purpose device for the reliable assessment of problem-solving ability. Its wider use is suggested on a trial basis. Such a trial will enable more detailed analyses of the instrument itself, and also its refinement. Further, the process by which the problem-solving assessment instrument was developed has potential for the development of other instruments to measure other generic skills. A suggestion is made that instruments to assess other key

competencies are developed using the development methodology that produced the problem-solving assessment instrument. The development of comparable instruments for other key competencies would allow the reporting of key competencies profiles which may prove to be attractive to graduates of the vocational education and training (VET) sector and their potential employers.

Key findings

In the main study the problem-solving assessment instrument was shown to be a reliable instrument for the assessment of problem-solving performance across a range of tasks within the Electronics and Information Technology program at Torrens Valley Institute of TAFE.

The instrument was also shown to work well in a validation study undertaken in the Certificate IV in Workplace Assessment and Training course, also at Torrens Valley Institute.

The assessment processes that were employed in trialling the problem-solving assessment instrument, which involved both self-assessment and lecturer validation, not only led to the assessment of the key competency of problem-solving, but also to its development among participants.

A strong relationship was found between problem-solving performance and educational achievement in the units of competency that students undertook.

Future directions

As a result of the research conducted in the Authentic Performance-based Assessment of Problem-Solving Project, several suggestions are made for the further development of the problem-solving assessment instrument and for the extension of the methods used in the project to other generic skills domains. These include:

- that more extensive trials of the problem-solving assessment instrument and its associated assessment processes be conducted in a wider range of courses and with a greater number of participants
- that the problem-solving assessment instrument be revised on the basis of feedback from further trials with diverse samples of learners
- that other key competencies assessment instruments be developed using the methodology that gave rise to the problem-solving assessment instrument
- that robust analytical techniques, such as those employed in this project, be used to ascertain the number of performance levels that can be discriminated for each of the key competencies that are to be assessed
- that forms of reporting based upon robust analytical procedures of reliable data be developed to better inform all stakeholders of the key competencies achievements of VET sector graduates.

1 Introduction

Background to the study

Following the release of the Mayer report (Mayer Committee 1992), there was much activity aimed at validating and implementing the key competencies (Curtis et al. 1996; Jasinski 1996). Although this research activity did demonstrate that these generic skills were accepted as being important, the extent to which they have been integrated into educational programs is disappointing. Recently, interest in both Australia and elsewhere in generic skills has been renewed (Curtis & McKenzie 2002; Kearns 2001).

Some of the problems, recognised by the Mayer Committee, that attend attempts to define generic skills, to implement them in education and training programs, and then to assess and report upon their attainment, remain. However, several recent changes in Australia's education and training systems now give reason to approach these problems with renewed confidence. In a report that presaged major changes to senior secondary assessment in New South Wales, McGaw (1997) recommended the use of criterion-referenced (standards-based) assessment and that the key competencies in all curriculum areas be identified specifically. Griffin (2001) has shown that criterion-referenced assessment can be used to gauge student achievement in higher order thinking skills and that a scale of performance based on this approach provides a foundation for reporting achievement in both the school and the vocational education and training (VET) sectors.

Assessing problem-solving

Problem-solving is consistently rated among the three most important of generic skills, the others being an ability to communicate and an ability to work well within teams (Field 2001). Generic skills are not applied in isolation: they are used in concert and their application is context dependent (Hager & Beckett 1999). However, efforts to assess and report on them that attempt to take a holistic approach may not address many of the subtle issues that specifically influence problem-solving performance and its measurement (Curtis 2001). An approach that acknowledges the complexity of problem-solving performance and its assessment is required.

Theories of problem-solving have existed for many decades and an evaluation of them has revealed a substantial convergence of views concerning the key constituents of problem-solving (Mayer 1992). Unfortunately, many efforts to assess problem-solving do not reveal an understanding of these perspectives. Problems that reflect a particular context are set and performance on them is graded (Herl et al. 1999). However, without a more generalised view of problem-solving, there is no assurance that the abilities that are assessed on a specific set of test problems will be applicable, and transferred to new situations.

The key problem-solving elements that have been recognised include problem identification, elaboration, strategy generation, enacting a chosen solution path and monitoring performance. However, problem-solving also varies with both the field in which it occurs and the level of expertise of the problem-solver in that field. Thus, in addition to general problem-solving strategies, both access to knowledge and the way that knowledge is used influence problem-solving performance, and these factors need to be included in a comprehensive approach to the assessment of problem-solving ability.

Developments in assessment and reporting

Assessment may have many purposes, but three central ones are to assist learning, to measure individual achievement and to evaluate programs (Pellegrino, Chudowsky & Glaser 2001). Assessing an aspect of students' learning signals the importance of that course component. Being able to measure individual problem-solving achievement and report on it in ways that are meaningful to both students and prospective employers should help to raise the profile of this important generic skill.

The renewed interest in criterion-referenced (standards-based) assessment that has occurred following the McGaw report (McGaw 1997) has stimulated research on performance-based assessment. Griffin and others have shown that it is feasible to employ this technique in the assessment of higher order thinking skills (Griffin 2000; Griffin et al. 2001). The study reported here has extended this to the integrated, authentic, performance-based assessment of problem-solving within VET sector courses.

Producing valid and reliable assessments depends upon having clear criteria across a range of tasks so that students have multiple opportunities to practise and demonstrate their emerging abilities. These criteria have been embedded in a new problem-solving assessment instrument, which was used in this project as the basis for scoring student performances. This instrument was used to assess students on a range of problem tasks that are part of the instruction and assessment practices of their courses.

A final element of assessment is the generation of a student score that provides a sound basis for drawing inferences about each student's ability. The output of the application of the problem-solving assessment instrument is a set of individual task scores. These were found to form a coherent scale of performance, and this coherence was tested using the Rasch measurement model (Bond & Fox 2001). In addition to establishing coherence, the technique also enabled the production of scaled scores that formed a sound basis for reporting (Masters & Forster 2000).

Purpose of the study

The purpose of this study, the Authentic Performance-based Assessment of Problem-Solving Project, was to demonstrate that key developments in cognitive theories of problem-solving and in assessment, measurement and reporting can form the basis of valid assessments of problem-solving performance.

This has been done using tasks that are embedded in current instructional and assessment practices of existing courses using the problem-solving assessment instrument to assess major problem-solving processes specifically.

A further purpose was to investigate the relationship between demonstrated problem-solving performance and learning within a course. It was hypothesised that students with higher levels of initial problem-solving strategies would learn at a faster rate within the course. Correlations between problem-solving ability as measured with the problem-solving assessment instrument and the grades achieved on the technical assessments in the course were taken as indicators of the relationship between problem-solving ability and educational achievement. Measures of approach to, and confidence in, problem-solving derived from the problem-solving inventory (PSI) (Heppner & Petersen 1982) were also used in examining this relationship.

The context

The project was conducted in the Electronics and Information Technology Program at the Tea Tree Gully campus of Torrens Valley Institute of TAFE in South Australia. Since its inception, the

Electronics and Information Technology Program has had a history of innovation in its delivery, and this context is outlined below.

Flexible learning

When students commence their Electronics and Information Technology Program courses at Torrens Valley Institute, they do so in class groups and are inducted into the program. They move from conventional class groups to a flexible learning mode when they feel confident to do so. In this mode, students enrol in as many units of competency or modules as they are able to complete given their other commitments. Instruction is provided through comprehensive support materials and in one-to-one and small-group sessions with the lecturer for that unit or module. Course materials and other resources are available in an open-plan laboratory and in other specialist facilities. These facilities are open from 9:00am to 9:30pm weekdays throughout the year and are also open on a limited basis on weekends. Students thus exercise considerable control over their study programs.

However, students are not simply left to themselves. Each student is assigned a mentor and student progress is monitored to ensure that they remain on track. Mentors help students plan their study programs and provide advice on any reorganisation of their programs as their commitments change.

Self-direction of students' study programs has been and remains a significant feature of the teaching and learning program in electronics and information technology. This is an important element of the context in which the Authentic Performance-based Assessment of Problem-Solving Project was conducted.

Current practice in assessment of key competencies

The Electronics and Information Technology Program at Torrens Valley Institute has included the study of generic skills since its establishment in 1990, before the release of the Mayer report. At that time, the skills that were developed were called enterprise skills, but were similar to the key competencies identified in the Mayer report (Mayer Committee 1992). Following the release of the Mayer report, key competencies were adopted to replace enterprise skills.

Since 2000, the approach adopted for the assessment of key competencies has been:

- Key competencies are promoted within the program, students are informed about them and about the availability of assessment in them, and students are encouraged to have them assessed.
- The assessment of key competencies has been embedded in the technical and vocational assessments that are specified in course modules and units of competency, that is, key competencies both have been embedded and are specifically assessed.

The flexible delivery of the Electronics and Information Technology Program has encouraged students to be self-directed. At a conference on key competencies in 1995 Eric Mayer made the point that key competencies were in effect a measure of students' self-directedness (Eric Mayer, personal communication). Thus, the delivery methodology of the Electronics and Information Technology Program and the implementation of key competencies were seen to be highly compatible.

The performance levels used in the assessment of key competencies have reflected the notion of self-directedness. For example, the performance levels for the key competency 'solving problems' are:

- Level 1: Solves the problem by following an established procedure.
- Level 2: Solves the problem by selecting from several alternative established procedures.

- Level 3. Creates a new procedure or adapts an existing procedure to meet the demands of a task.

For a student to be credited with a key competency at a particular level, it must be demonstrated on at least two occasions on two different tasks. This is in accordance with the guidelines proposed in the Mayer report.

An implication of the adoption of the notion of self-directedness as a key element differentiating the levels of performance is that higher levels do not necessarily subsume lower ones. For example, it may be possible for a student to be creative in developing a novel solution strategy but not to be as proficient in following set procedures. For this reason, students are encouraged to seek key competency assessments at all levels.

Because of the history and context of the Electronics and Information Technology Program at Torrens Valley Institute, the view may be taken that the results of the study will not generalise to other contexts. However, it should be noted that key competencies assessment was voluntary, that the rate of uptake was not as high as the teaching team had hoped and that this project has been an impetus for including greater numbers of students in key competencies assessment. Further, in parallel with the main Authentic Performance-based Assessment of Problem-Solving Project investigation, a validation study was conducted in a program, also at Torrens Valley Institute, that does not share the history of involvement in key competencies or the self-directed flexible learning context of the Electronics and Information Technology Program. Nonetheless, the extent to which the findings can be generalised to other contexts remains an issue for further investigation.

Terminology

Many terms have been used to describe the sorts of skills that have been named key competencies in Australia, essential skills in New Zealand, essential skills and employability skills in Canada, necessary skills (for the workforce) in the United States, and key, or core, skills in the United Kingdom. Many other labels have also been applied. They include soft, generic, transferable, and even life, skills.

A substantial volume of policy work has been undertaken recently by three major employer and industry organisations in Australia: the Australian Industry Group, the Australian Chamber of Commerce and Industry and the Business Council of Australia (Allen Consulting Group 1999; Australian Chamber of Commerce and Industry & Business Council of Australia 2002). As a result of this work there is considerable debate about what generic skills should be called and what constructs the term should subsume. In order to discuss these constructs in an inclusive way, the term key competencies is used in this report when referring specifically to the set of skills identified in the Mayer report (Mayer Committee 1992). The term generic skills is used to refer to the broad collections of these skills. This term, generic skills, is probably one of the least appealing of the alternatives, but it is also one of the most general and is used for that reason.

2 Literature review

This chapter begins with a review of the development and implementation of key competencies in Australia. This covers both the period leading up to the release of the Mayer report and the implementation phase that followed. Also reviewed are some recent industry reports that have sought to re-energise debate on this matter. These developments establish the importance of problem-solving as a necessary generic skill and place it in a context provided by other skills that are also regarded as significant.

A review of approaches to assessment is presented. This section examines both developments in assessment practices in general and approaches that have been taken to the assessment of generic skills. This section seeks to establish a context for the development of the approach to the assessment of problem-solving that has informed the Authentic Performance-based Assessment of Problem-Solving Project.

Theories of problem-solving are analysed in order to establish a sound and coherent definition of this construct and its components and a basis for its assessment that has both construct and content validity.

The chapter concludes with a statement of the two research questions that directed the research efforts that are reported in this document.

Key competencies in Australia

The development of key competencies

Three major reports that led to the definition and description of key competencies in Australia are generally recognised to be the Karmel (Quality of Education Review Committee 1985), Finn (Finn Review Committee 1991) and Mayer (Mayer Committee 1992) reports. These developments are summarised below. The Carmichael report is also significant for the contribution it made in establishing the structural framework for vocational education and training (Employment and Skills Formation Council 1992).

The Karmel report

The terms of reference for the Karmel report (Quality of Education Review Committee 1985, pp.204–205) recognised the growing importance of an internationally competitive labour force and Karmel recommended that specific outcomes of education should be subject to monitoring and reporting in order to ensure that Australia's education systems contributed to the nation's competitiveness (recommendation 10, p.203). Karmel recommended that achievement in basic skills and in certain generic skill areas, including mathematics, science and technology, should be assessed. These areas were later recognised by the Finn and Mayer committees as significant. The focus of the Quality of Education Review Committee was on enhancing outcomes of compulsory education as preparation for both further education and work.

The Finn Review

The Finn Review (Finn Review Committee 1991) was asked, among other wide-ranging terms of reference, to report on 'appropriate national curriculum principles designed to enable all young people ... to develop key competencies' (p.2).

The Committee drew attention to changes in the skill demands of industry and to rapid change in the Australian economy as a result of international competition and structural economic change nationally. It noted that:

> the most successful forms of work organisation are those which encourage people to be multi-skilled, creative and adaptable. (Finn Review Committee 1991, p.6)

The Committee argued that, because of changing technologies and economic circumstances, 'the ability to continue learning and acquiring new or higher level skills will be fundamental'. As a consequence 'the emphasis of our training system has to be both on the acquisition of the specific skills for the job/trade and on flexibility', and flexibility 'requires a strong grounding in generic, transferable skills' (Finn Review Committee 1991, p.55).

The Committee further noted a recognition by employers that students required:

> a foundation of basic skills and a range of broad skills and attributes which are generally relevant to the world of work without being occupation- or industry-specific. (Finn Review Committee 1991, p.6)

The Committee recommended that emphasis be given to six key areas of competence: language and communication, mathematics, scientific and technological understanding, cultural understanding, problem-solving, and personal and interpersonal skills (Finn Review Committee 1991, p.58).

The Committee also recommended that an expert group be established to undertake more detailed work on defining and assessing the initial list of proposed key competencies. The work required of that group was:

- ✧ to elaborate the basic concept of key competencies
- ✧ to operationalise it for the school and training sectors
- ✧ to specify levels of achievement
- ✧ to recommend arrangements for assessing and reporting on student achievement.

That group was chaired by Eric Mayer and reported in 1992.

The Mayer report

In order to establish the scope and definition of generic skills, the Mayer Committee reviewed developments overseas (especially in the United Kingdom and the United States), consulted with industry and with educators in the school and VET sectors and to a lesser extent in the higher education sector, and finally undertook a validation exercise which involved further consultations with industry.

The extensive involvement of the school and VET sectors reflected a concern at the time with post-compulsory education and training, mainly for 15 to 19-year-olds, and with the pathways available to them in moving from compulsory education to employment or further study.

The Mayer Committee accepted the National Training Board's definition of competence:

> The concept of competence adopted by the National Training Board includes these elements: 'it embodies the ability to transfer and apply skills and knowledge to new situations and environments'. This is a broad concept of competency in that all aspects of work performance, not only narrow task skills, are included. (Mayer Committee 1992, p.7, citing National Training Board 1991)

The requirements of key competencies were defined by the Mayer Committee as follows:

> Key Competencies are competencies essential for effective participation in the emerging patterns of work and work organisation. They focus on the capacity to apply knowledge and skills in an integrated way in work situations. Key Competencies are generic in that they apply to work generally rather than being specific to work in particular occupations or industries. This characteristic means that the Key Competencies are not only essential for participation in work, but are also essential for effective participation in further education and in adult life more generally. (Mayer Committee 1992, p.7)

The Committee summarised their requirements for key competencies by saying that they must:

- be essential to preparation for employment
- be generic to the kinds of work and work organisation emerging in the range of occupations at entry levels within industry, rather than occupation or industry specific
- equip individuals to participate effectively in a wide range of social settings, including workplaces and adult life more generally
- involve the application of knowledge and skill
- be able to be learned
- be amenable to credible assessment.

One of the key areas of competence recommended by the Finn Review Committee, cultural understanding, was discussed by the Mayer Committee, but eventually was not included as a key competency.

The key competencies that were eventually endorsed were:

- collecting, analysing and organising information
- communicating ideas and information
- planning and organising activities
- working with others and in teams
- using mathematical ideas and techniques
- solving problems
- using technology

Implementation of key competencies

Considerable efforts were expended in embedding the key competencies in the school and VET sectors and to a lesser extent in the higher education sector. There were substantial pressures for change in both the school and VET sectors during the early 1990s—the period during which the key competencies were to be implemented. The changing context of the school sector has been well-documented (Lokan 1997).

Change in the VET sector

The VET sector has undergone very considerable change over the past decade. The major changes have included the establishment of the Australian National Training Authority (ANTA) in 1995, greater national co-ordination of the sector, curriculum change and new course delivery strategies, and greater and more direct involvement of industry in specification of training requirements.

A trend towards curricula specified in terms of outcomes to be achieved has been more marked in vocational education and training than in other sectors of education. At a time when ministers for education and training were seeking ways of prescribing outcomes from the sector, competency-

based training was endorsed. This fits well with industry's requirements that people emerge from training programs with the skills that are required on the job. Industry-specific skill requirements are specified in training packages that emerge from consultative processes mediated through industry training advisory bodies.

A further and major change in the VET sector during the 1990s has been the creation of an open training market with both state-funded and private providers and in which the principle of user choice is embedded.

It was within these substantial structural changes in vocational education and training that the key competencies were introduced into the sector. Considerable effort was expended, especially in the mid-1990s, in implementing key competencies in the sector. However, the scope and pace of change have caused some implementation difficulties. In particular, the concept of generic competencies appears to have become confused with vocational competencies (Down 2000), the latter being industry specific and necessarily narrow in focus, while the former are meant to be very broadly applicable.

Key competencies in vocational education and training

Jasinski (1996) found that there was a diverse understanding of key competencies within technical and further education. This was portrayed positively as representing the different manifestations of key competencies in different training areas, but it may also have indicated a lack of conceptual clarity in the definition of key competencies. Jasinski advocated an extension of the scope of the key competencies to include 'entrepreneurialism, learning competencies, and intra-personal competencies' (Jasinski 1996). Reynolds and van Eyk (1996) reported that among non-TAFE VET providers there was little understanding of key competencies and that the term competency, used in both vocationally specific and generic senses, created confusion. Down (2000) also found that some confusion arose between industry-specific competencies and key competencies. She reported that key competencies have not been prominent in training packages in the past, and that they were seen by some practitioners as desirable but optional components. At that time, support materials for the developers of training packages did not include advice on the implementation of key competencies—a matter that has since been rectified. The more recently developed training packages, for example, the Business Services Training Package, have paid specific attention to key competencies.

Other research conducted within TAFE (for example, Curtis 1996; Lawson & Hopkins 1996; Russell 1996) found evidence that the key competencies were recognised and accepted as valuable, but they also reported that some understandings of the generic nature of the skills and of ways in which they might be embedded within training programs were limited.

Jasinski (1996) reported that there was little support within the VET sector for the use of the proposed performance levels for the key competencies. Down (2000) found that the assessment levels proposed by the Mayer Committee for the key competencies were confused with the levels of the Australian Qualifications Framework.

However, Keeves and Kotte (1996) demonstrated that the measurement of the key competencies was possible. They suggested that the performance levels proposed by the Mayer Committee could be assigned greater meaning by using the Biggs and Collis (1982) Structure of the Observed Learning Outcome (SOLO) taxonomy as an organising framework. Keeves and Kotte (1996) concluded that:

> Research and development studies in these three areas of competence—mathematics, science and technology, and language—have the capacity to transform in a meaningful way the movement towards certification in terms of key competencies that the Mayer Committee has proposed. (Keeves & Kotte 1996, p.116)

This proposition provides a basis for an expectation that meaningful measurement and reporting can be extended to other generic competencies.

Experience with generic skills in workplace training

Generic skills are recognised as being important for individuals, for enterprises and for industry (Hase 2000; O'Keefe 2000). Hase (2000) described the importance of teamwork, creativity, learning to learn and self-efficacy in the development of individual and organisational capabilities. The conception of capability described by Hase reinforces the importance of key competencies as developed by the Mayer Committee (Mayer Committee 1992), but also suggests that they need to be extended if high-performance workplaces are to develop more broadly. Indeed, new frameworks for representing generic skills continue to emerge (see, for example, Australian Chamber of Commerce and Industry & Business Council of Australia 2002; Kearns 2001). The key competencies described by Mayer do incorporate those skills recognised as being the most important, but the newer schemes emerging in Australia and elsewhere include a wider range of attributes.

Analysis of the impact of generic skills on workplace learning needs to recognise that many enterprise-based training programs are not necessarily constructed around the competencies derived from Mayer. While much workplace learning is based on industry-specified training packages, in some enterprises it is not, but rather is structured to meet enterprise-specific needs. Such material often has a strong emphasis on various forms of generic competence, but they are not necessarily expressed in the terminology used by Mayer (Field 2001) and they embrace skills that are not included in the Mayer key competencies. Dawe (2001) also found that the generic skills embedded in training packages did not completely coincide with the Mayer key competencies. Young workers seem to develop their generic skills through a range of interactions: they develop some generic skills through formal instruction, some through work experience and some through their interactions with their families and social networks of more experienced peers (Smith 2000).

The findings reported above, in particular the inconsistent understandings of key competencies, their varied representation in training packages, the use of other generic skill elements, and the various ways in which generic skills are learned by different learners, all suggest that the Mayer-specified key competencies are not being consistently promoted, presented or assessed in the VET sector in Australia. However, key competencies are still regarded as being important, and several major industry bodies have sought to refocus debate about these constructs and to re-energise their implementation in all sectors of education and training in Australia.

Recent Australian generic skills initiatives

Underpinning a recent resurgence of interest in generic skills in Australia and in other developed economies is a view that increasing global competition and the growing use of information and communication technologies will propel these economies towards 'knowledge economy' industries and higher levels of performance. In order to support these trends, a highly skilled workforce is required. However, it is not sufficient to have high levels of industry-specific technical skills, although they are required. Continuing change in advanced economies will require highly flexible and adaptable workforces, and so a mix of skills that includes basic, or foundation, skills, highly developed technical skills and well-developed generic skills will be required of the workforce as a whole and of individuals. The generic skills component provides a basis for the flexibility and adaptability that characterise workforces that are capable of evolving to meet the changing requirements of a developing economic order. The contextual factors influencing Australia's skill requirements are summarised in a recent report by the Australian Chamber of Commerce and Industry and the Business Council of Australia (2002, pp.7–8).

There are consequences for individuals as well as for the workforce as a whole. It is an oversimplification to assert that because Australia's workforce is experiencing change all employees must adapt to new roles: many existing jobs will remain. However, the experience of the United States indicates that 'old jobs', especially low-skill jobs, will experience a decline in income and that high-skill jobs will attract increases in salaries, creating a growing gap between old and new jobs and low and high-skill jobs (Bailey 1997). There is a clear public policy concern to create the workforce that will enable Australia to participate in the new economy, but there is also a public policy concern to minimise the number of individuals who are caught up in the low-skill, low-wage scenario and who are at risk of declining employment opportunities and greater reliance on welfare support. Thus policy initiatives must address the national perspective at a system level and must also tackle skill development at the individual level.

Training to compete

In 1999, the Australian Industry Group commissioned a report on the training needs of Australia's industries (Allen Consulting Group 1999). Among many findings, the report noted that:

> an increasing premium is being placed on generic skills, both 'hard' (notably IT [information technology] skills) and 'soft' (e.g. problem-solving, team skills, willingness and ability to adapt) to be developed prior to recruitment. (Allen Consulting Group 1999, p.v)

The report then outlined the skills that are required by Australian industry if it is to remain globally competitive (p.xi). These included, in addition to job-related technical skills, sets of core, or basic, skills, interpersonal skills and personal attributes. These are summarised in table 1.

Table 1: Summary of the sets of generic skills proposed by the Australian Industry Group

Generic core, or basic, skills	Interpersonal, or relationship, skills	Personal attributes
Literacy	Communication	Capacity to learn
Numeracy	Team working	Willingness to embrace change
Information technology capability	Customer focus	Independent problem-solving and reasoning capability
Understanding of systems relationships	Project and personal management	Practicality and a business orientation

Source: Allen Consulting Group (1999, p.xi)

The Australian Industry Group classification of skills is an interesting departure from the list proposed by the Mayer Committee. First, it includes reference to basic, or core, skills. This has been shown to be a problem area for Australia and for other developed economies. The International Adult Literacy Survey has shown that a significant proportion of Australia's adults have levels of literacy and numeracy that are inadequate to function effectively in the current work situation (OECD & Statistics Canada 1995). Indeed, the assumption of adequate basic skills made by the Mayer Committee was challenged in its consultations (Mayer Committee 1992, p.95). Second, the Australian Industry Group classification separates interpersonal and personal skills, and subsumes some of the traditional 'intellectual' skills, such as reasoning and problem-solving, under personal attributes. The main focus of the Mayer Committee was on the set of intellectual skills.

Employability skills for the future

More recently, the Australian Chamber of Commerce and Industry and the Business Council of Australia, with support from the Australian National Training Authority and the Commonwealth Department of Education, Science and Training, undertook a comprehensive study of generic employability skills in Australia and elsewhere (Australian Chamber of Commerce and Industry & Business Council of Australia 2002). Their methodology included an extensive literature review of the Australian and overseas situations, the conduct of focus groups and interviews with individuals

from small, medium and large enterprises, and a validation exercise involving extensive consultation with companies and employer organisations.

The report reiterated many of the previously identified contextual factors that are expected to continue to influence Australia's competitiveness. These include a demand for greater profitability, increasing global competition, and increased complexity, innovation and flexibility (Australian Chamber of Commerce and Industry & Business Council of Australia 2002, p.25).

The report proposed an employability skills framework, and recognised the importance of the Mayer key competencies as a basis for continuing work in this area. However, the report also identified the need to expand the scope of employability skills to include personal attributes and supported a more extensive list of skills than had been recognised in earlier work (Australian Chamber of Commerce and Industry & Business Council of Australia 2002, pp.26–27). The major skills groups identified are (pp.29–38):

- communication
- teamwork
- problem-solving
- initiative and enterprise
- planning and organising
- self-management
- learning
- technology

Important workplace abilities, such as customer service and leadership, result from combinations of elements of the major skills. The employability skills framework provides a coherent list of attributes and skills while also providing scope for flexibility at the enterprise level. The framework is summarised in table 2.

Several issues arising from the employability skills framework are noteworthy. First, it does include extensive lists of skills and attributes and it has certainly broadened the scope of the employability skills concept beyond the Mayer Committee's approach. Second, each of the skills has been elaborated through lists of skill 'elements', and these provide an opportunity for each skill to be contextualised. Thus, the framework acknowledges the variation in skill requirements of different work contexts, while retaining the central concept of broadly applicable generic skills. Third, the Australian Chamber of Commerce and Industry and Business Council of Australia report is clear that some work-related skills are in reality combinations of more central key skills. For example, customer service involves both communication and problem-solving. An important implication of this is that it is not necessary to develop an exhaustive list of skills; it is more productive to identify a common set of skills that, in combination, lead to high job-related performance.

The framework proposed represents a new approach to the definition of generic skills, compared with the Mayer Committee's recommendations. The Mayer Committee specifically excluded values and attitudes. These are prominent among the personal attributes of the employability skills framework. The eight skills groups in the framework include both traditional intellectual skills and personal and interpersonal skills. These skills groups are extensively elaborated in the report (Australian Chamber of Commerce and Industry & Business Council of Australia 2002, pp.26–38).

Table 2: Summary of the Australian Chamber of Commerce and Industry/ Business Council of Australia employability skills framework

Personal attributes		
Loyalty	Personal presentation	Balanced attitude to work and home life
Commitment	Commonsense	
Honesty and integrity	Positive self-esteem	Ability to deal with pressure
Enthusiasm	Sense of humour	Motivation
Reliability		Adaptability
Key skills		
Communication skills	that contribute to productive and harmonious relations between employees and customers	
Teamwork skills	that contribute to productive working relationships and outcomes	
Problem-solving skills	that contribute to productive outcomes	
Initiative and enterprise skills	that contribute to innovative outcomes	
Planning and organising skills	that contribute to long-term and short-term strategic planning	
Self-management skills	that contribute to employee satisfaction and growth	
Learning skills	that contribute to ongoing improvement and expansion in employee and company operations and outcomes	
Technology skills	that contribute to effective execution of tasks	

Source: Australian Chamber of Commerce and Industry & Business Council of Australia (2002, pp.26–27)

Generic skills in Australia: Summary

The release of the Mayer report was a watershed in Australia's education and training systems. For the first time, a requirement to develop, assess and report upon generic outcomes of education and training programs at both individual and system levels was endorsed. The principal focus of the specified key competencies was employability, but the Committee also noted their significance for individuals in their participation in Australian society.

The Mayer Committee restricted the list of endorsed key competencies to skills that were teachable and assessable and specifically excluded values and attitudes. During the Mayer Committee's extensive consultations with industry and community groups, there had been requests to include a broader range of skills and abilities. However, the Committee restricted the scope of its recommended competencies in accordance with the principles for key competencies that it had developed.

Recent developments in generic skills in Australia—specifically the Australian Industry Group, and Australian Chamber of Commerce and Industry and Business Council of Australia reports—suggest that the definition of generic skills needs to be broadened in order to meet the needs of Australia's emerging economic context. These developments are consistent with trends reported in other countries (Curtis & McKenzie 2002; Kearns 2001). The changed definition of generic skills poses challenges for the development, assessment and reporting of achievement of these skills. Whether it is feasible to extend the Mayer approach to a broader conception of generic skills will depend upon a demonstration that the proposed new generic skills meet some of the criteria specified by the Mayer Committee—that they can be delivered, assessed and reported upon credibly.

The assessment of key competencies

Section 1.5 of the Mayer report, 'Assessing and reporting achievement of the key competencies' (Mayer Committee 1992, pp.41–56), deals extensively with both assessment and reporting issues. It recommends nationally consistent assessment and reporting of individual achievement of the key competencies (p.42). It also recommends that generic skills be assessed on several occasions in

different contexts, thus recognising the need to demonstrate the generic nature of these abilities. For each of the key competencies, the Mayer Committee recommended that three levels of performance be recognised. However, the report does not specify precisely how these abilities should be assessed, and this matter has been the subject of many subsequent investigations. These are summarised below, in the section 'Current approaches to the assessment of generic skills'.

The Committee then moved on to reporting issues and recommended reporting at the individual level through a 'record of performance' using a common format (p.51). The Committee also recommended reporting at an aggregate national level that was to be based upon statistical sampling of individual records of achievement, and that the performance of equity and target groups should be a specific focus of this approach (p.55). Individual and aggregated reporting, however, have distinct purposes which may require different assessment approaches.

Before pursuing the Mayer position and its implications, it is worth examining some recent developments in assessment practices.

Assessment

Assessment is rightly one of the very strongly contested areas of educational theory and practice. Assessment has a range of purposes and many different methods are employed, each with a particular set of characteristics. It forms a basis for reporting individual achievement and can be used to evaluate system-level performance. It is often a 'high-stakes' activity with very significant consequences for individuals. Assessment is also important for other stakeholders. When an education provider asserts that an individual has attained a specified level of achievement, employers and the community can reasonably expect that judgement of performance to be dependable. In order to meet these expectations, education providers must themselves ensure that their practices and processes meet acceptable standards.

In addition to developments in assessment, there have been related developments in measurement and reporting. These are highly relevant to the assessment and reporting of achievement of key competencies and are reviewed below.

Purposes of assessment

Airasian (1994) and Pellegrino, Chudowsky and Glaser (2001) asserted that assessment has three broad purposes:

- to assist learning
- to measure individual achievement
- to evaluate programs.

Appropriate assessment at the individual level may lead to enhanced individual learning, in part by signalling that what is being assessed is regarded as being important. The signalling function of assessment may be particularly important in situations where the assessment is mandated by agencies outside the learner–instructor interface, since importance is indicated both to learners and to their teachers. A corollary of this is that where assessment of particular attributes is not mandated, low importance is being signified.

Assessment also reveals individuals' achievement and this may be useful information to both the individuals and their potential employers, indicating areas of strength and weakness.

Aggregation of individual achievement can be used at the system level to monitor system performance. However, great care must be taken in the sampling design and analytical phases of this approach to ensure that biased performance estimates are avoided.

Thus, while there are different purposes for assessment, appropriate assessments at the individual level can be combined to monitor program, institutional and system-level achievement. However, what constitutes appropriate assessment for a given purpose is contested.

Validity and reliability of assessment

The appropriateness of an assessment may be judged on its validity and its reliability, and many dimensions of validity have been described.

Validity refers to 'the adequacy, appropriateness, and usefulness of inferences that can be made on the basis of test scores' (American Educational Research Association, American Psychological Association & National Council on Measurement in Education 1985). Zeller (1997) identified several types of validity—construct validity, content validity, concurrent validity and predictive validity—each of which must be demonstrated in order to show that a measure is valid. Osterlind (1998) argued that each of the types of validity requires that different forms of evidence be assembled in order to show overall validity. Gillis and Bateman (1999) reviewed validity and reliability in the context of competency-based assessment in the VET sector in Australia.

Construct validity requires that a case be made to show that the conception that underlies the proposed assessment is itself coherent and that descriptions of the concept lead directly to the form and content of the assessment activities. This has particular relevance to the assessment of key competencies. Although key competencies have been described and are 'generally understood', in order to provide a sound basis for reporting achievement, they must be defined with sufficient precision to enable specific indicators to be established.

Content validity requires that the breadth of the concept that is to be assessed is completely and fairly reflected in the assessments that are proposed. In order to do this, the construct that is being assessed must be described in complete detail and the importance of each component of the construct identified so that it can be adequately represented in the assessment tasks.

Evidence for *concurrent validity* emerges from a comparison of the results of the proposed assessment with other outcomes that are expected to produce similar results. This might be demonstrated by employing an alternative form of assessment that is designed to measure the same construct. However, the obverse of concurrent validity is *discriminant validity*, and to demonstrate this it is necessary to show that the construct and its proposed assessment are unique in some important respect. This concept is also important in considering key competencies. A range of 'mapping exercises' have been undertaken to show that certain key competencies are embedded within particular units of technical competency. On this basis, the attainment of a relevant technical competency is taken as *de facto* evidence of the achievement of the key competency, but this fails the test of discriminant validity.

Predictive, or criterion, validity requires that the assessment results correlate with measures of other constructs that are thought to be contingent upon the concept being assessed. In the case of key competencies, it is rather difficult to do this directly as the claimed benefit of high achievement of key competencies is enhanced workplace performance. Indirect evidence for this relationship is likely to be found by comparing the personnel of claimed high performance with those of other workplaces.

Reliability is demonstrated by showing that the assessment produces a consistent result for a person of a given ability, over time and irrespective of who administers or grades the assessment. Different forms of assessment require different methods for demonstrating reliability. Where judgements of performance are made, consistency of judgement, or inter-rater reliability, must be demonstrated. For paper-based tests, a range of questions is tested in calibration studies, and subsets of questions that yield common and consistent results emerge. These provide the basis for alternative forms of a common assessment instrument.

In addition to validity and reliability, the precision of the assessment is also an important consideration, and this requires genuine measurement of performance.

Measurement

In the past, the scores derived from the administration of assessment instruments were assumed to provide the 'measure' of student performance. The definition of measurement given by Stevens in 1946, that 'measurement is the assignment of numerals to objects or events according to a rule' has been used as the basis of measurement claims in the social sciences since (Michell 1997). Michell showed that Stevens' definition was a necessary but insufficient basis for true measurement. It does not support 'additivity' of scores, as scores typically obtained through this process of rule-based assignment produce rank scores and not interval ones (Harwell & Gatti 2001).

Where scoring tools are to be used as the basis for assessing and reporting student achievement, it is essential that they be based upon valid and reliable conceptions of the construct being reported and that they are capable of generating estimates of students' achievements with sufficient precision to justify the purposes for which the assessments are to be used. Where fine distinctions are to be drawn among levels of performance and where high-stakes assessments are used, very high precision instruments are required. The precision of instruments must be known and it must be sufficient for the intended purpose.

Authentic assessment

Much school assessment has been pencil-and-paper based and these forms of assessment have been criticised as being inauthentic. This criticism is advanced against large-scale standardised testing because it rewards decontextualised atomistic knowledge and high levels of verbal fluency. Therefore, this form of assessment is regarded as being invalid, as verbal ability rather than, or in addition to, the target ability is being tested (Wiggins 1989).

As a result of this criticism, other forms of assessment have been sought and many alternatives have been tried. Many of these involve judgements made by raters of student performances on 'real-world' tasks. The final judgement is influenced by student ability, as it should be, since this is what is intended to be assessed. However, even though the tasks may be authentic, the judgement is also influenced by the characteristics of the task on which the learner is being assessed and by an interaction between the learner and the task. No doubt some tasks offer greater opportunity to demonstrate a skill than do others, but there may be some tasks about which some individuals have greater confidence than others and this is likely to be reflected in the assessed performance. Similarly, where a range of judges is being used, some students may encounter an 'easy' judge while others may encounter a severe one. On the same task, students of a particular ability level could achieve quite different marks. Again, there is a task difficulty factor, a judge severity factor, and an interaction between judges and the tasks they are asked to rate. These factors—task difficulty, rater severity, and interactions between students and tasks and raters and tasks—all contribute to variation in the final judgements of performance and reduce their precision.

In the VET sector in Australia, authentic learning has been identified as having its own problems. Robertson et al. (2000) noted that, although workplace learning occurred in an authentic environment, the learning outcomes were of variable quality. They noted in particular that key competencies might not be addressed in these situations.

The sources of variation that influence a judgement of final performance have been reviewed. Inter-rater reliability is often a concern in large-scale assessments and indices of this are frequently reported in research papers. Shavelson, Gao and Baxter (1993) examined a database of performance assessments and took a 'sampling framework' view of each of the contributors to variation in performance judgement. They concluded:

> that rater-sampling variability is not an issue: raters (e.g., teachers, job incumbents) can be trained to consistently judge performance on complex tasks. Rather, task-sampling variability is the major source of measurement error. Large numbers of tasks are needed to get a reliable measure of mathematics and science achievement at the elementary level, or to get a reliable measure of job performance in the military. (Shavelson, Gao & Baxter 1993, p.iii)

This has implications for the assessment of both technical competence and key competencies. The challenge is to devise tools that enable the task-related contributions to variability of performance judgement to be reduced to acceptable levels. In turn this will increase the precision of the measures that are derived from the assessments to levels appropriate to the interpretations that arise from them.

Levels of assessment

An important assessment issue is the choice between single benchmarks and multiple levels of performance. In competency-based training, it is common to specify a single benchmark level of acceptable performance which is either achieved or not achieved. Given the breadth of application of these important skills, the diversity of learners and the diversity of industries and occupations into which they will move, simple benchmarks of performance are likely to be insufficient. There is little support among employers for a single benchmark, especially if it is pitched at a minimum competency level. In addition, in assessing complex performance, such benchmarks may have counterproductive influences on the achievements of cohorts of individuals, as their training providers may feel compelled to ensure that the maximum number of individuals achieve the benchmark rather than encouraging each person to achieve their individual maximum (see Masters & Forster 2000). Assessment at several levels is one way of ensuring that each individual strives to achieve a maximum and therefore contributes to high levels of workplace performance.

Summary of assessment issues

The issues of validity, reliability and precision are significant in establishing levels of measurement. The Mayer Committee recommended three performance levels for each of the key competencies (Mayer Committee 1992, pp.18–19). Although three levels were described, there is an implicit level 0 which is identified when level 1 is not achieved. The Committee acknowledged that the number of levels and their specification should be subjected to review.

In part, the levels described by the Mayer Committee reflect increasing management by the individual of the processes that underlie each of the competencies. In this sense, the described levels reflect a degree of metacognitive control over the component processes and move beyond simply describing the performance as a behaviour. The levels identified by the Mayer Committee are quite appropriately criterion referenced, but they are more than simply behavioural descriptions of performances, and this additional dimension might usefully be reflected in instruments developed to assess performances.

Whether three levels can be discriminated depends on the precision with which judgements can be made. If the precision is high, more levels become discriminable, but if the precision is low, it may not be possible to distinguish even three performance levels. The precision of the measures derived from assessments must be both known and sufficient to meet their intended interpretations.

The number of levels that are required depends upon the purposes of identifying levels and interpretations that might be based on them. If the purpose is feedback to the individual about their developing abilities, three levels may be adequate. If the purpose is to assess the performances of cohorts of employees in the context of establishing benchmarks for the requirements of high-performance industries, additional levels may be necessary.

Current approaches to the assessment of generic skills

The framework established through consideration of the validity, reliability, authenticity and precision of assessments is used to compare a range of approaches to the assessment of generic skills. Four broad approaches to the assessment of generic skills have been identified in a review of the relevant literature:

- holistic judgements by teachers (McCurry & Bryce 1997; National Industry Education Forum 2000)
- portfolios created by students (Feast 2000; National Industry Education Forum 2000; Reynolds 1996)
- assessment based on work experience (National Industry Education Forum 2000; Queensland Department of Education 1997)
- assessment using purpose-developed (standardised) instruments (Australian Council for Educational Research 2001b; Griffin 2000; Herl et al. 1999).

These approaches are not competing alternatives. They achieve similar purposes—to document and certify student achievement—through different means and, because of their relative strengths, they complement each other.

Holistic judgements

Of the four approaches listed above, teacher judgement has been shown to work well in the school sector, where teachers know students' attributes well through frequent and close observation (McCurry & Bryce 1997, 2000). McCurry and Bryce were able to establish small panels of teachers and to provide sufficient training in the key competencies to enable them to make consistent judgements of students' attainment of key competencies. This training and the observation of students, in both classroom-based and co-curricular activities, enabled teachers to make sufficiently consistent judgements to discriminate eight performance levels.

However, this method is unlikely to transfer to either the VET or the higher education sector, where such close observation does not occur. A trainer may monitor learners in an institutional classroom or workshop setting, but is unlikely also to see them in social or workplace settings. Similarly, workplace supervisors may observe individuals in that context, but not in others. The bases of judgements made by these raters will be restricted to a limited range of contexts. The conclusions reached by Shavelson et al. (1993) suggest that it may not be possible to generalise from their limited observational contexts.

Portfolio assessment

Portfolios may be quite effective for making students aware of their developing skills and for providing a rich data source for detailed scrutiny by prospective employers. Several encouraging examples of portfolio-based assessment of generic skills are available (Conference Board of Canada 2000; Reynolds 1996; Troper & Smith 1997). Devices such as the 'Employability skills toolkit for the self-managing learner' (Conference Board of Canada 2000) have an important instructional function in that they reveal to learners the key dimensions of generic skills and they provide a framework in which learners can document their achievements and present supportive evidence.

One of the disadvantages of this form of assessment is that, while it provides a very detailed account of a learner's achievements, it is not in a form that is readily digestible or comparable. The Michigan Workforce Readiness Portfolios scheme described by Troper and Smith (1997) sought to address this problem by training raters to assess the portfolios, and a summary assessment was presented with each portfolio. However, Troper and Smith expressed some doubts about the

validity and reliability of portfolio assessment. Validity is compromised because a very good portfolio may well represent a person with highly developed generic skills, but it also reflects a person who is capable of assembling a persuasive document, and this is a separate, albeit important, skill. Reliability depends upon having the portfolios assessed by independent and trained raters. Troper and Smith suggested that portfolios be used for low-stakes, but not for high-stakes, purposes. However, portfolios are used in hiring decisions, and for the individual, this is a high-stakes application.

Workplace assessment

Assessment based on work experience appears to be a useful method and to produce a simple report, although, like portfolio assessment, it is not standardised and may not be amenable to ready comparisons. The National Industry Education Forum (2000) approach to portfolio assessment combines teacher judgement, self and peer assessment, and workplace assessment to produce a comprehensive but standard-format portfolio. The comments made in relation to the assessment of portfolios produced by students are also applicable to this form of assessment. Indeed, workplace assessment might be considered to contribute an important component to a student's overall portfolio.

The comments made by Robertson et al. (2000) are apposite here. The quality of both the learning and the assessment depends on the context provided by the workplace and the knowledge and diligence of the assessors. In addition, in relation to Shavelson et al.'s (1993) notion of generalisability, variations in workplace contexts will contribute to variation in opportunities to develop generic skills and therefore to a lack of precision in the results and in inferences that can be drawn from them.

Standardised instrumental assessment

Independent assessment using standardised and purpose-developed instruments enables efficient assessment and provides a basis for reporting that is readily interpreted by learners and potential employers. However, the criticisms of standardised testing, outlined in the section 'Authentic assessment' above, suggest that this approach suffers low validity, even though its reliability and precision are high.

One criticism of independent assessment using standardised instruments is that it decouples assessment from teaching. However, individuals learn in a great variety of situations, including schools, TAFE colleges, universities and workplaces, yet all need to demonstrate similar sorts of competencies at different times. Thus, specifying the outcomes, rather than attempting to specify both learning processes and outcomes, leaves some flexibility for learners and their teachers. This is analogous to the situation in relation to training packages, in which learning objectives and assessment processes are stipulated, but in which the curriculum through which the skills are developed is a matter of judgement for the training provider, and a matter of competitive differentiation among providers in an open training market.

Two instrumental approaches to the assessment of generic skills are reviewed here: the graduate skills assessment (Australian Council for Educational Research 2001a) and the Center for Research on Evaluation, Standards and Student Testing (CRESST) report on problem-solving assessment (Hedl et al. 1993).

The Graduate Skills Assessment Project

The Graduate Skills Assessment Project (Australian Council for Educational Research 2001a) is a pilot study to ascertain whether it is feasible to assess and report on the generic skills of university

graduates. There have been several administrations of this instrument during the life of the project and approximately 3600 students from 28 universities and all major fields of study have participated. It has been administered to students at or near entry to undergraduate courses and to other students on exit.

The instrument has tested problem-solving, critical thinking and reasoning and written communication. Claims can be made for its concurrent validity as scores on the test correlate with course entry scores (tertiary entrance ranks). Data analysis has shown that the instrument is reliable and that it has good discriminations among the components, supporting claims for content validity. However, because it is a pencil-and-paper test, it is an easy target for the criticism that it lacks authenticity.

There are some informative findings from this project. It has shown clear differences among major discipline groups on the overall test score and on the component scores. For example, medical students have done better than humanities students (most likely reflecting admission criteria). Engineering students have fared better than nursing students on the problem-solving component, while nursing students have performed better than engineering students on interpersonal skills (Hambur & Glickman 2001). While it may be possible to criticise the instrument for favouring some groups because of the test format, the identification of different generic skills profiles for different course cohorts is expected. Even though the skills being tested are truly generic—that is, they are required across all represented disciplines—particular disciplines require or reward the skills differentially. This has implications for the assessment of key competencies. A common pattern of performance on all key competencies should not be expected for all industry groups. If different profiles are expected, then identifying these differences will require assessment and analysis, leading to measurement of sufficient precision to permit the necessary distinctions to be made.

An interesting feature of the Graduate Skills Assessment Project has been the assessment of students at or near entry and near graduation from their courses. For individuals, this will enable a 'gain score' to be reported, but this score is much more informative when it is aggregated, using appropriate statistical methods, at the course, institutional and system levels. Appropriate multilevel analyses of data collected at the individual level will enable comparisons of the 'value added' through particular courses and institutions. Consistent with one of the Mayer objectives, this form of analysis could identify the extent of 'value adding' for equity target groups.

At this stage, the graduate skills assessment test has limited scope, but consideration is being given to extending it to include basic skills, management skills, information technology skills and research skills.

The assessment methods that have been used in the Graduate Skills Assessment Project have demonstrated that it is feasible to assess important components of employability skills validly. The construction of the assessment instruments and the methods of analysis used (the Rasch measurement model) have enabled measures of students' performances to be reported along a scale, and for informative interpretative comments to be provided for users. Additional analyses of those performances will lead to the establishment of differentiated levels of performance. Further, a result of the Graduate Skills Assessment Project has been the identification of the distribution of students' achievement. On this basis, individuals can receive a score on each competency, along with mean scores for all students and for students in their particular course. The set of scores on each of the competencies assessed constitutes a generic skills profile for each student. Such a profile should be a useful document for prospective employers, as they can see the individual's performance and compare it to national means and, once industries have become accustomed to these reports, to industry expectations and needs.

The Graduate Skills Assessment Project provides a model for an instrument that could be developed specifically for the VET sector, and this could provide a criterion-based assessment measure to complement other, perhaps more authentic, forms of assessment that are undertaken within courses and workplaces.

The Centre for Research on Evaluation, Standards and Student Testing (CRESST) model of problem-solving assessment

The second instrumental model that is worthy of consideration is that trialled by the Center for Research on Evaluation, Standards and Student Testing (Herl et al. 1999). This approach sought to standardise an authentic assessment approach. Two tasks were selected for problem-solving assessment. One involved a bicycle pump and the other the human respiratory system. Both tasks were conducted as pencil-and-paper activities, but they did not use the conventional multiple-choice format. Diagrams of the system (for example, the bicycle pump) were presented and students were asked a series of questions about what would happen under certain conditions. They were also asked to explain what fault might lead to certain described symptoms. Answers to these questions were assessed using a scoring rubric. Of interest in this project was the use of concept mapping. Students were asked to label diagrams with a range of pre-prepared labels that represented concepts and relationships. Clearly, there was an attempt to avoid some of the criticisms of most pencil-and-paper tests. While the novel use of concept mapping was an interesting aspect of the test, it could also have added a new dimension. Students who were familiar with this technique would have been at an advantage and would have scored higher than students, with equal problem-solving ability, without this background. The other aspect of the test that gives cause for concern is the narrow range of tasks. Given the comments of Shavelson et al. (1993) about the sampled nature of tasks, the generalisability of this assessment approach is open to question. In addition, this approach, although pencil-and-paper based, was not efficient for either the participants or their assessors.

A future for generic skills assessment

Considerable work remains to be done on the assessment of generic skills. Additional elements ought be included in the assessment of generic skills in order to demonstrate that the full range of these skills can be assessed reliably. In addition, the assessment needs to be compatible across the school, VET and higher education sectors. There are some complexities in this expansion. The issue of levels of performance becomes more complex, although having a broader range of participants may make the identification of levels of performance more accurate and certainly it would make it more useful. In this respect it is worthy of note that the early trials of the graduate skills assessment instruments included a substantial number of upper secondary students (Australian Council for Educational Research 2000).

Some of the elements of generic skills that are more difficult to measure include personal and interpersonal skills. A number of instruments through which these skills are assessed exist (Mayer 2001; Salovey et al. 1995). The measurement of attitude is an important component of the personal dimensions of generic skills and methods in this field are well-established (see, for example, Anderson 1997; Wright & Masters 1981). These methods can be applied to the assessment of generic skills.

Given the range of purposes that have been identified for generic skills assessment, it seems that several approaches to assessment will be required. The main characteristics of assessment approaches are that, collectively, they should provide:

- a mechanism for communicating the scope of generic skills to learners, training providers and employers
- a means of providing feedback to learners on their acquisition of generic skills and a framework for their improvement
- a rich source of information about individual achievement, with supportive evidence
- an opportunity to undertake assessments that are authentic and occur within a work context or one that closely simulates it

- a method of assessment that is not onerous for either the learner or the assessor
- a summary of the performance of individuals that is readily accessible by employers
- a cost-effective means of collecting performance information, individually and at aggregate (institutional and system) levels.

The suite of assessment and associated reporting arrangements described above, namely teacher judgement, portfolio assessment, workplace-based assessment and instrumental assessment, collectively meet most of the desirable criteria for generic skills assessment. Each has been shown to be effective in particular situations. What remains to be shown is that they can be used in concert to meet all objectives for the assessment and reporting of generic skills at both individual and aggregate levels.

Problem-solving

Problem-solving has been recognised as an important and desirable attribute in generic skills schemes described in Canada, the United States, the United Kingdom, New Zealand and Australia. Recently, in industry consultations undertaken by Field (2001), employers reported that problem-solving was among the three most valued generic skills. However, despite its widespread acceptance and its perceived value, descriptions of it are rather imprecise. In order to develop sound assessments, it is necessary to develop a detailed description of problem-solving. This is now done through a review of major theoretical perspectives of problem-solving.

A definition of problem-solving

Mayer (1992, p.1) presents a useful definition of problem-solving:

> Problem solving is cognitive processing directed at achieving a goal when no solution method is obvious to the problem solver.

The key elements of this definition are, first, that it involves cognitive processes. These are not directly observable, but must be inferred from the actions of the problem-solver. Second, it requires that the action be goal directed. This goal may be specified in the problem scenario or it may be formulated as a result of initial inquiry by the problem-solver. Third, the definition requires that a solution method not be immediately obvious or available to the problem-solver and therefore that the problem-solver must search for a solution method, either through the recall of related problems that have been encountered previously or through the application of generalised problem-solving heuristics. An important implication of this definition is that certain activities that may be regarded by some as problems, for example, some mathematics tasks that involve the application of drilled procedures, should not be accepted as problems and therefore should not be accepted as providing evidence of problem-solving proficiency.

The definition of problem-solving presents a challenge to the performance levels for problem-solving developed by the Mayer Committee. The Committee recommended that three levels of performance be recognised for all key competencies.

> Performance Level 1 describes the competence needed to undertake activities efficiently and with sufficient self management to meet the explicit requirements of the activity and to make judgements about quality of outcome against established criteria.
>
> Performance Level 2 describes the competence needed to manage activities requiring the selection and integration of a number of elements, and to select from established criteria to judge quality of process and outcome.
>
> Performance Level 3 describes the competence needed to evaluate and reshape processes, to establish and use principles in order to determine appropriate ways of approaching activities, and to establish criteria for judging quality of process and outcome.
>
> (Mayer Committee 1992, p.18)

Mayer performance level 1 is not consistent with the definition of problem-solving, since an individual at this level is required to undertake activities according to explicit requirements, whereas the psychological definition requires a search for an appropriate solution method. However, this performance level does require monitoring of the solution process, and this is a characteristic of problem-solving. The remaining two performance levels do conform to increasing levels of problem-solving difficulty and reflect increasing requirements for performance monitoring.

Theoretical perspectives of problem-solving

Mayer (1992, p.i) opens his preface with the following quotation from Wertheimer (1959):

> Why is it that some people, when they are faced with problems, get clever ideas, make inventions and discoveries? What happens, what are the processes that lead to such solutions?

The mixture of successes and failures among the many competing theories of problem-solving suggests that there are places for general-purpose strategies and conditioned knowledge structures and for higher order processes that manage the application of knowledge, skill and strategy in seeking solutions. In attempting to assess problem-solving, general-purpose cognitive processes, evidence for the activation of knowledge, and the role of the learning context must be included.

Greeno, Collins and Resnick (1996) outline three major theoretical stances on learning and problem-solving that they term associationist/behaviourist/empiricist, cognitive/rationalist, and pragmatist/situative/sociohistoric. Of these, the variants of the second and third, cognitive and situative, stances appear to be most useful in developing understandings of individuals' development of problem-solving capability.

Among cognitive approaches to problem-solving, different aspects of the problem-solving process are emphasised. The information processing approach (Newell & Simon 1972) is one cognitive approach, in which the problem-solver is perceived as a general-purpose information processing entity who apprehends the problem in a task environment and forms an abstract representation of it to construct the problem space. The problem-solver then searches the problem space in order to find pathways from the current state to the goal state.

A key element of the information processing approach is that the problem-solver uses the most general and abstract information processes both to represent and to solve the problem. A deficiency of information processing as a theory of human problem-solving is that many subjects fail to transfer general processes from one form of a problem to analogous forms. For example, in one study subjects did not transfer the procedure for solving the traditional disks-on-pegs version of the Tower of Hanoi task to abstract analogues, such as monsters and globes (Kotovsky, Hayes & Simon 1985).

An alternative cognitive model arises from comparisons of novice and expert approaches to problem-solving. On the basis of their experience, experts have built highly developed domain-specific knowledge structures called schemata. Experts use their extensive knowledge base and their deep understanding of the principles of the domain, whether it be physics, computer programming or playing chess, to form meaningful representations of the problem situation, and then use automated processes to move towards solutions. Novices, who lack the well-developed knowledge structures of the domain, tend to be influenced by the surface features of problems and must rely upon less-efficient general-purpose problem-solving processes that make greater demands on working memory.

Comparisons of experts and novices have led to a focus on the role of schemata in effective problem-solving. Schemata are cognitive structures in which elements of knowledge, of several types, are organised into chunks. In experts, these chunks of related knowledge are activated as a problem is apprehended, and they play a role in the representation of the problem, in planning a solution method and in the execution of the solution.

The mainly cognitive views of problem-solving have been criticised by researchers who have investigated problem-solving in 'real-world' work and social situations. For example, Scribner (1986) described the processes by which packers assembled orders in a dairy warehouse, and Lave (1988) showed how shoppers made decisions about which products, sold in different quantities, represented best buys. These individuals showed reasonably high skill levels in the domains in which they operated and what the researchers noted was that they did not use the abstract arithmetic processes that they had learned at school. Instead, they manipulated concrete materials, used non-formal practical representations, and developed and used low mental load strategies. However, the problems studied were all relatively simple, well-structured problems, with obvious goals.

Mayer (1992, pp.506–507) quoted Rogoff (1984):

> Thinking is intrinsically woven with the context of the problem to be solved ... Our ability to control and orchestrate cognitive skills is not an abstract context-free competence which may be easily transferred across widely diverse problem domains but consists rather of cognitive activity tied specifically to context.

Supporters of the situated cognition approach argue that a purely cognitive model of problem-solving that does not address the context in which problems are presented cannot adequately describe the problem-solving process. However, each model has provided useful approaches to understanding problem-solving and the implications of each need to be considered in developing problem-solving assessments.

Implications of theories of problem-solving for assessment

Among the cognitive approaches to problem-solving, two alternative conceptions have been presented: one emphasises general problem-solving processes and the other stresses the role of highly developed domain-specific knowledge structures. There is little doubt that each is important for people in different situations. One of the purposes of vocational education and training is to develop in learners expertise in the domain of their employment or, at the very least, the capacity to develop that expertise in practice.

Although characteristic differences between novices and experts have been observed in many fields, the processes by which novices become experts have not been well-explained. But that is precisely the purpose of education and training programs—to develop in people the capacity to become more expert in their fields. Problem-solving is the set of processes that people use to deal with novel situations, to use what resources they have, and to reflect upon and learn from their solution attempts.

In the context of vocational education and training, the situated or social cognitive approach has apparent relevance. Scribner's dairy packers were expert at what they did, and one of the goals of vocational education and training is to pass on or develop that sort of expertise. However, the problems that were studied were well-defined, had obvious goals and were 'low-stakes' activities. It may be a matter of small moment if a shopper purchases a small container of peanuts instead of a larger one that might be a little cheaper per gram. In work situations, maximising productivity will require optimal solutions to problems, and so they tend to be high-stakes activities. These high-stakes activities also represent an end point in the application of knowledge. Thus they have been the target of transfer and there may be little point in using them as indicators of performance in other domains. In assessing problem-solving, the search is for people who demonstrate that they have skills in one domain that they are likely to use when confronted with novel problems. Tasks in which problem-solving skill is demonstrated thus become source, not target, tasks for transfer.

An important implication of the situative approach to understanding problem-solving relates to the context in which problem-solving occurs and is assessed. The situative model suggests that valid assessment is likely to occur only in contexts that are like the intended contexts of application and

that pencil-and-paper tests of this ability are unlikely to provide useful indications of likely performance in work-related situations.

Thus a major question remains: What are the processes by which novices transform themselves into experts and what would an observer look for in seeking evidence of problem-solving ability? An answer to this question requires an analysis of the problem-solving processes that have been described.

Cognitive processes in problem-solving

A search for the most general cognitive processes reveals both performance components and metacognitive components. Sternberg (1985) lists selective encoding, selective combination and selective comparison as performance components, while Mayer (1992) labels similar processes as selecting, organising and integrating. Metacognition includes having an awareness of cognitive activities, monitoring them and regulating them.

Many authors have contributed to the body of knowledge that describes problem-solving processes. Polya (1957) described four major processes in problem-solving: understand the problem; devise a plan; carry out the plan; look back. Bransford and Stein (1984, p.12) listed five processes: identify problems; define and represent them with precision; explore possible strategies; act on these strategies; look back and evaluate the effects on your activities. Mayer and Wittrock (1996, p.50) included: assess the requirements of the problem; construct a solution plan; select an appropriate solution strategy; monitor progress towards the goal; modify the solution plan when necessary. These are but a few of the many descriptions of problem-solving processes that are present in the very extensive literature on this topic.

Common among the descriptions of problem-solving processes are:

- apprehending that a problem exists and elaborating or defining the problem
- planning an approach to the problem including selecting strategies
- carrying out the plan
- monitoring progress towards the goal
- looking back at the progress of the solution attempt.

These sets of processes may be combined to form fewer categories or split to form more, and various authors have presented them in a variety of combinations and have emphasised different skill sets. Apprehending that a problem exists and defining it and execution of the chosen plan are performance processes. Planning, monitoring and reflection are all metacognitive activities, although they involve performance components.

In the scoring scheme that has been developed for this project (see 'Development of the problem-solving assessment instrument' in chapter 3), greater emphasis is placed on metacognitive processes as these are both the most general and are thought to be the ones that are most implicated in transfer.

However, these cognitive and metacognitive processes operate on, or within, knowledge structures, and the role of problem-solvers' knowledge must also be included in an account of problem-solving and in its assessment.

Key factors in successful problem-solving

There are many other elements of problems and problem-solvers that should be considered in the development of a scheme for scoring student performance on problems. The characteristics of the problem tasks would include whether the problems presented:

- were well or ill-defined
- depended only upon the application of known knowledge (knowledge transfer) or upon the application of specific or general problem-solving methods (problem-solving transfer)
- required the development of novel solutions or approaches
- were sufficiently complex to demand decomposition into separate sub-problems. (This has implications for 'unit-of-analysis' considerations.)

These matters will affect the difficulty of the problem task and should be reflected in the scores that are achieved on particular tasks.

The characteristics of problem-solvers themselves also influence performance. Prior knowledge and interest in the topic are both known to influence success (Alexander & Judy 1988; Murphy & Alexander 2002). Problem-solvers' beliefs about their own abilities, both in general terms and within the domain of the problem tasks that they face, also influence their performance, both directly and through their motivation to persist with difficult tasks (Bandura 1997). In the present study, these characteristics were not assessed in detail. The problem-solving inventory (Heppner & Petersen 1982) was used to assess three affective elements of students' approaches to problem-solving: problem-solving confidence, approach–avoidance style and personal control.

The assessment of problem-solving

For an assessment of problem-solving to be useful, it must be based on a conceptually sound framework and have the attributes of validity that depend upon such a framework. The assessment must have face validity, and thus be recognisably an assessment of the construct described by theory. In the previous section, several perspectives on problem-solving were outlined.

The review of literature has suggested that characteristics of the problem-solver, the problem environment and aspects of the problem task all influence problem-solving performance. The problem-solver brings to the situation a combination of relevant knowledge, a set of general problem-solving processes, and a set of affective factors including motivation. The environment may influence individuals' performances, but it is not a variable that has been controlled in this study. A range of tasks was selected as productive opportunities for students to demonstrate their abilities. Some tasks were thought to be reasonably simple and should have enabled participants to experience success while others were more challenging. However, students' overall performance on the task as a whole was not tested in the problem-solving assessment. The focus of the assessment was on students' demonstrated problem-solving processes. Even so, those tasks that were the most challenging were likely to enable students to demonstrate higher levels of problem-solving ability than easier ones, and it is reasonable to expect that the tasks selected for this assessment would influence the level of performance demonstrated. The variable of interest is the learner's problem-solving ability, but observation of this is likely to be influenced by the task and the person making a judgement. Shavelson et al. (1993) concluded that the problem task is a greater source of variability than raters, although it must be noted that past assessments of problem-solving have focussed on performance in the domain rather than on an evaluation of the processes that are employed by the individual.

The approach to the assessment of problem-solving used in this project is based upon the hypothesis that problem-solving performance that is likely to be broadly applicable across tasks and domains comprises a core set of processes that are enacted when an individual is confronted with a novel situation. Thus, this is not a test of the type of problem-solving that experts would follow within their domain of expertise. It is a test of the processes that people can call upon when faced with, and through which they learn to deal with, novelty and uncertainty.

Research questions

Two main questions have driven this research.

- Is it feasible to extend existing problem-solving tasks through the development of scoring rubrics that address problem-solving processes in order to develop reliable assessments of problem-solving ability?
- Is there an identifiable positive relationship between indicators of problem-solving ability and subsequent learning and use of knowledge about problem-solving tasks?

These questions carry with them many implicit assumptions, and these are teased out through the methods adopted and the analyses of the data that were collected.

3 Methods

The Authentic Performance-based Assessment of Problem-Solving Project is part of an ongoing research program to investigate the definition, assessment and reporting of generic competencies in Australia.

Project development and approvals

The Authentic Performance-based Assessment of Problem-Solving Project was developed as a collaborative venture between the Centre For Lifelong Learning and Development and Torrens Valley Institute of TAFE, Adelaide, and it forms part of longer term research and development and continuous quality improvement programs in these organisations.[1]

Because of affiliations of the two collaborating organisations with the Flinders University of South Australia and the then South Australian Department of Education, Training and Employment (DETE), approval was sought and gained from the Social and Behavioural Research Ethics Committee of Flinders University and the Strategic Planning and Information: Research Unit of DETE. The research has been conducted in accordance with the requirements of the joint National Health and Medical Research Council and Australian Vice Chancellors Committee Statement and Guidelines on Research Practice (Australian Vice Chancellors Committee 1997). Specific requirements of this code include:

- Common good: The research was proposed because the knowledge generated as an outcome of the study should result in improvements in the reporting of educational achievement and in the processes employed within courses that lead to the desired outcomes—the enhancement of generic competences.
- Harm minimisation: This project was such that no harm could occur to participants as a result of their involvement. Indeed, it could only help participants develop more informed conceptions of their problem-solving ability.
- Informed consent: Participants, both individuals and the institution, have been informed of the purposes, processes and expected outcomes of the research.
- Confidentiality: Individual participants are guaranteed that the results of the study will not lead to their identification in any published reports arising from the study.

Both staff and students were informed of the purposes, processes and anticipated outcomes of the project. Staff were consulted through forums held with members of the Electronics and Information Technology Program teaching team and other members of the Institute through meetings of the Torrens Valley Institute key competencies focus group. Students were provided with extensive information about the project. Consultation occurred with the Electronics and Information Technology Student Representative Council; information was communicated through posters that were displayed prominently in areas frequented by students; screen savers were used in student

1. This project is also a pilot study within a research program being undertaken by David Curtis as part of his PhD program at the Flinders University of South Australia.

laboratories to promote the project; email messages were circulated to students; and the intranet that students used to access course resources was used to disseminate information. In addition, students were given detailed information about the project in the consent forms that were distributed (see appendix 1). The project was launched at a luncheon attended by industry representatives who were invited to speak about the importance of generic skills in the workplace.

Security and confidentiality of completed information forms and assessment instruments have been assured through their secure storage. Participants' names and other identifying information such as the assessed tasks, for both staff and students, have been removed from data files and replaced with ID codes.

Recruitment and characteristics of participants

The project was designed to be conducted in the Electronics and Information Technology School at Torrens Valley Institute of TAFE. Courses range from Certificate II to advanced diploma levels. Students are quite diverse in age and in education and work backgrounds. Some students study full time, but most study part time and have work commitments ranging from casual and part time to full time.

All students undertaking electronics and information technology programs were informed of the project through the processes outlined in the previous section and were invited to participate. In addition to these promotional efforts, staff noted students' progress through their modules and, when students were about to undertake a module that included one of the assignment tasks that had been identified as having potential for the demonstration of problem-solving ability, an individual email was sent to them advising them of the opportunity available. However, participation was voluntary and no pressure was exerted on students to participate.

Thirty-three students participated in the main study and completed research consent forms. Some other students also submitted technical assignment tasks for problem-solving assessment, but did not complete research consent forms, so their data have not been included in this report.

In addition to the main study, a subsequent validation study was undertaken, also at Torrens Valley Institute of TAFE, involving 48 learners who were enrolled in the Certificate IV in Assessment and Workplace Training.

Selection and development of instruments

Two instruments were used in this research project. The problem-solving assessment was the principal instrument. It is a modification of a prototype developed to assess problem-solving skills. In addition, in order to assess participants' attitudes towards problem-solving, the problem-solving inventory (Heppner & Petersen 1982) was used. The development and administration of these instruments are now described.

The problem-solving inventory

The problem-solving inventory (PSI) was developed by Heppner and Peterson (1982). In its original form it had 35 items, three of which were 'test' items. Six response categories were used. These categories were neither labelled nor described in the original paper. The problem-solving inventory has three subscales: problem-solving confidence (11 items), approach–avoidance style (16 items) and personal control (5 items). The factor loadings, obtained in exploratory factor analysis reported by its developers, were above 0.4 except for four items, all from the approach–avoidance style subscale.

The instrument was shown not to measure the same thing as either intelligence tests or social desirability. It showed concurrent validity with other measures of problem-solving and locus of control. Cronbach alphas for the three subscales were 0.85, 0.84 and 0.72, and for the whole instrument, 0.90. Split-half and test–retest reliabilities were 0.85, 0.88 and 0.83, and 0.89 for the whole instrument (Heppner & Petersen 1982).

The instrument has been reviewed favourably by Camp (1992), who found that, although primarily a research instrument, it could also be used for 'contrasting problem-solving appraisals with more objective measures of actual abilities' (p.699).

In order to avoid response-set effects, the original form of the instrument included 15 reverse-scored items. Some of the items included double negative statements. For example, the first item read:

> When a solution to a problem was unsuccessful, I do not examine why it didn't work.

Osterlind (1998) suggested that approximately 10–15 per cent of items should be reverse-scored in order to detect response-set effects. The almost 50 per cent of reversed items in the instrument was thought to detract from its readability, and so statements with double negatives were reworded. In the revised instrument 11 reverse-scored items remained, but only five of these included negative statements.

It was believed that the six unlabelled response categories were both unnecessary and unclear. These were reduced to four categories and they were labelled 'almost always', 'often', 'seldom' and 'almost never'. The revised version of the problem-solving inventory is shown in appendix 3.

Development of the problem-solving assessment instrument

As outlined in the review of literature, most past efforts to assess problem-solving have focussed on individuals' performance on selected problem-solving tasks. However, the tasks chosen have been shown to contribute a substantial component of performance variability and therefore to mask the contribution of individual ability to performance variability. Since the purpose of the problem-solving assessment is to identify individual ability, approaches in which this is contaminated by other factors are compromised.

In past efforts to assess problem-solving in a componential, rather than a holistic, way, separate scoring rubrics were developed for each task (Herl et al. 1999; Shavelson et al. 1993). If this approach were to be taken in the VET context in Australia, the load on assessors would be excessive. Each training package has many units of competency and each unit has many tasks. The process of developing separate rubrics for this number of tasks and then of providing professional development to ensure that they were used consistently would be onerous at system and provider levels and for individual assessors. Thus, in this project an intention was to develop either a single instrument or a very small number of generally applicable instruments.

The problem-solving assessment instrument was designed to assess the use of problem-solving processes directly, as these processes are thought to be important in the emergence of expertise within a domain and also to be transferable between tasks within and possibly between domains. The development of the instrument followed four stages:

- the identification of a coherent theoretically sound construct
- the identification of major component processes
- the identification of indicators of those processes
- the establishment of levels of performance on each indicator.

The stages in the development of the problem-solving assessment instrument are now described.

A coherent conception of problem-solving

In chapter 2, several major approaches to describing problem-solving were canvassed (see the section Problem-solving). Its applicability to the assessment of problem-solving within domains for non-experts was the major factor in selecting the chosen approach, which focusses upon the application of general problem-solving processes, rather than upon access to and use of domain-specific knowledge. In addition, considerable agreement was found among many authors on the types of general processes that are implicated in problem-solving. A set of these processes is now outlined.

Major problem-solving processes

The processes that are central to effective problem-solving were identified as:

- apprehending, identifying and defining the problem
- planning an approach to the problem including selecting strategies
- carrying out the chosen plan
- monitoring progress towards the goal
- reflecting on the effectiveness of the solution attempt.

In the problem-solving assessment instrument (see appendix 4), these processes have been labelled representation, planning, execution, monitoring and reflection.

The identification of indicators of problem-solving processes

For each of the five major processes a set of indicators was sought. Indicators of performance are the basic elements of measurement—they are the items that form the hypothesised scales of the problem-solving construct. In the case of the problem-solving assessment instrument an overall scale of problem-solving ability is hypothesised, that is, it is assumed that there is such an overall scale of problem-solving ability and that all the indicators form a single factor that represents that construct. Further, the single factor does have components, each of which is internally coherent and contributes to the single factor. In measurement terms, it is hypothesised that there is a set of subscales that together contribute to a coherent overall scale.

Each measurement scale must reflect fully the content domain of the concept, and so each component process must be represented in the overall scale (content validity). However, to be practically useful, the scale must have a limited number of items—probably between 15 and 25—and so each component process was limited to between three and six indicators. The indicators that were finally selected are shown in the instrument (see appendix 4). One indicator was added to the execution process: application of strategies. It was added to reflect the three performance levels specified in the Mayer report. In the analysis of data, this indicator was used to test whether the three levels originally proposed for the 'solving problems' key competency were supported.

The indicators used in the problem-solving assessment instrument were established by answering the question 'What would a competent person do in showing that they were able to apply the component process in a real situation?'

The establishment of performance levels for indicators

Several bases for the establishment of performance levels are available. The performance levels specified in the Mayer report could have been used. However, they were designed to be pragmatic, and there is no apparent theoretically sound basis for their use. They have been useful in making holistic judgements of performance, but finer grained judgements have been shown to be reliable (McCurry & Bryce 1997).

An alternative is a revised version of Bloom's Taxonomy of Cognitive Objectives (Anderson & Krathwohl 2001). This provides six levels of cognitive skill: remember, understand, apply, analyse, evaluate and create. While these cognitive skill descriptions do represent a sequence of increasing cognitive skill, they were designed to operate in conjunction with the levels of the knowledge dimension of the revised taxonomy, and do not appear to be sufficiently consistent with the major problem-solving processes that were derived from the literature search.

A second alternative, the Structure of the Observed Learning Outcome (SOLO) taxonomy (Biggs & Collis 1982), was investigated. This taxonomy is based upon the cognitive complexity of individuals' responses to the application of knowledge in learning and problem situations. It recognises levels of performance from ineffective use of knowledge to very complex and abstract application. A description for each level of the SOLO taxonomy is shown in table 3.

Table 3: Performance levels of indicators using the SOLO taxonomy

SOLO level	Description	Score
Pre-structural	No knowledge, inaccurate recall or does not use relevant knowledge	0
Uni-structural	Uses relevant knowledge/skill elements in isolation	1
Multi-structural	Uses relevant knowledge/skill elements in combination	2
Relational	Can generalise using knowledge within the problem situation	3
Extended abstract	Can extend what has been found through the problem to other situations	4

Source: Biggs and Collis (1982)
Note: SOLO = Structure of the Observed Learning Outcome

An advantage of the SOLO taxonomy is that its five levels form a set of ordered responses. These responses are thus amenable to analysis using item response theory and may form the basis of linear interval scales. This possibility is to be tested with data collected using the problem-solving assessment instrument, but if it is shown that item response theory can be used to score student problem-solving performance, a powerful tool for investigating the components of performance becomes available. This can provide information on the precision of assessments of problem-solving ability that in turn will indicate the number of performance levels that can be discriminated reliably.

The levels of the SOLO taxonomy have been applied to descriptions of performance on each indicator in the problem-solving assessment instrument and provide a means of scoring student performance. Not all SOLO levels have been applied to all indicators. In order to make the instrument as easy to use as possible, the number of SOLO levels selected for each indicator was based on the anticipated ability of assessors to make reliable judgements of student performance. Thus, for some indicators; for example 'sets a realistic goal', only two levels are suggested, while for others; for example 'plans an approach to the problem', four levels are available.

In the analysis of data collected as a result of this project, it is expected that some of the indicators may be revised or even deleted and that the number of performance levels for indicators may be revised.

Administration of the problem-solving assessment

The procedure for assessing students' problem-solving performances is outlined in appendix 2. The key steps in the procedure were:

- The student completes the initial project documentation (consent form, personal details, the problem-solving inventory).
- The student selects and undertakes the technical assessment task that is to be presented for problem-solving assessment.
- The student undertakes a self-assessment of problem-solving ability (see below).

- The student presents their technical assessment task and their self-assessment for validation by the facilitator.
- The student and facilitator discuss the technical assessment task and the student's self-assessment of their problem-solving ability.
- The facilitator records the result for later reporting and certification.

The assessment of problem-solving was based on an assignment that had been presented for technical assessment. The grade for this assessment was recorded. Then the same assessment activity was re-presented immediately for problem-solving assessment. Re-presenting the activity for problem-solving assessment immediately after the technical assessment was regarded as necessary to ensuring that the assessment of problem-solving was authentic.

Two key elements of the process are self-assessment by learners in order to collate evidence for the claimed problem-solving performance level and validation by the lecturer (facilitator). These are now described.

Self-assessment

The assessment of problem-solving in this project was intended to form an iterative process through which students learned what was meant by problem-solving so that they could improve the process in successive assessments. Thus the assessment procedure was intended to be a very overt process of both learning and assessment.

Students and staff used the same form of the problem-solving assessment instrument. This was done in order to ensure that the process was a completely open one in which all assessment criteria were clearly laid out for the students. They were aware of exactly how their lecturer would conduct that part of the assessment and what evidence they would be seeking. Indeed, the reverse side of the problem-solving assessment instrument included scoring instructions and sets of questions that related to each indicator used in the assessment. The questions were designed to focus attention on the evidence that might be expected to support the levels of performance that had been suggested for that indicator.

Through their initial self-assessment, students were expected to learn about the processes of problem-solving and how they are applied in authentic practice situations. They were also expected to become consciously aware of the processes as they applied them and to identify evidence of their application in practice. Thus, in subsequent assessment activities, students should become more proficient users of these processes and be better able to identify and present evidence of them in practice.

Intended outcomes of this process were that students would be both more proficient in their use of problem-solving processes and better able to describe what problem-solving is in practice and how they were able to implement problem-solving processes in specific cases. Students were expected to gain explicit knowledge of problem-solving, and therefore be able to describe it, and to gain explicit knowledge of how the processes are used in practice.

Lecturer (facilitator) validation

Although the term lecturer assessment was used in the early stages of the project, the term lecturer validation was adopted subsequently as a more accurate description of the process.

Lecturer validation comprised two elements. First, students were required to present their self-assessment forms and to present or point to evidence to support their self-assessment of problem-solving. The second element of the process was the lecturer's judgement of the student's problem-solving performance based upon the evidence presented by the student.

Once the lecturer had made their judgement, the assessment was discussed by the lecturer and the student. The purpose of this discussion was to draw attention to aspects of the assessment in which the student had either not presented evidence or had misinterpreted the evidence or the criteria. The discussion was clearly instructional, with the aim of enhancing the student's understanding of both problem-solving and the process of assembling evidence against specified criteria.

Reporting and certification

The performance level that was recorded in the student record system was determined by the lecturer. Since the assessment of key competencies at Torrens Valley Institute results in the award of a statement of attainment, accountability for the recorded level of attainment lies with the institution and its staff. It is essential that the processes that underlie a statement of attainment are credible and robust and that they would, if required, withstand an external review such as a quality audit.

Before the Authentic Performance-based Assessment of Problem-Solving Project was established at Torrens Valley Institute, key competencies had been assessed and reported and had resulted in the award of statements of attainment. Past practice assessed problem-solving using the three performance levels specified in the Mayer report. In order to provide continuity with past practice, and current practice in relation to the reporting of other key competencies, an indicator with three performance levels, corresponding to the Mayer performance levels, was added to the problem-solving assessment instrument. This indicator is 'application of strategies'. This item was used to establish the performance level recorded in the student records system (SMART). This item also provided an opportunity to examine whether the remaining indicators in the instrument supported three performance levels, a matter that is addressed in the analysis of problem-solving assessment data.

In addition to recording the results of the assessment, comments were recorded in the student records system. The comments field provided an opportunity to record details of the assessment task that was the vehicle for demonstrating the key competency and of aspects of the student's approach to the task. Thus the field provided a source of qualitative data for the project.

Data collection and analysis

The project accumulated data from four main sources. First, when students were recruited into the project they completed a consent form and a brief personal details form through which demographic data were acquired. At that time, students also completed the problem-solving inventory, which provided information on their approaches to problems, their confidence in problem-solving and their sense of control over problem-solving processes. The completion of these forms took approximately 20 minutes. The third, and principal, source of data for the project was the problem-solving assessment. The fourth data source was qualitative, and included comments made by students about the assessment process.

Quantitative data

The data that respondents recorded on printed versions of the instruments were entered into data files for subsequent analysis using a range of software packages including SPSS (SPSS Inc. 2000), Quest (Adams & Khoo 1999) and RUMM (Sheridan, Andrich & Luo 1997). Several approaches were taken to the analysis of quantitative data. In order to inform those who are familiar with more traditional approaches to data analysis, factor analyses and scale reliabilities analyses were conducted. However, one of the purposes for conducting the project, and for designing the instruments in the way that was chosen, was to establish valid measurements of the identified

constructs. In the discussion of measurement in the review of literature on assessment, the point was made that simply assigning numbers as a result of a set of scoring rules does not constitute measurement. A powerful tool for the conversion of ordered numerical scores to genuine verifiable measures of known precision is available through the Rasch measurement model. This is now described.

The Rasch measurement model

When tests or attitude survey instruments are administered, participants are scored according to whether they get items right or wrong in tests or according to the options they choose in response to attitude items. The raw numbers that result from this form of scoring indicate that, on a given item, a person achieving a higher score performs better than a person with a lower score. However, it is well-known that some test items are more difficult than others, and a correct response to a difficult item indicates a higher ability than a correct response to an easy item. The raw scores that result from tests can be misleading. Comparing two people who scored 10 and 11 on a test shows that one person scored one more mark than the other and has a slightly higher ability. However, comparing two others, one who scored 24 and the other 25 on a test out of 25, may suggest that the second person has a slightly higher ability than the first. But gaining that final mark may have been based on a very difficult item than should indicate that the person scoring 25 has a very high ability. In short, conventional scoring schemes do not produce interval scales. True measurement requires such scales.

The Rasch approach to measurement models responses to items and takes into account both the difficulty of items and the overall abilities of the respondents. The Rasch model is a statistical one that calculates the probability of a person of a given overall ability responding correctly to an item of known difficulty. High-ability people are very likely to score well on easy and moderately difficult items, but less likely to do well on very difficult items. Low-ability people are likely to do poorly on more difficult items. By modelling responses in this way, the Rasch method is able to use raw data and to assign levels of difficulty to items and estimates of ability to respondents. It is also able to detect aberrant or unexpected responses that are seen when low-ability students get difficult items correct by guessing or when items are unclear and are misunderstood by both high and low-ability students.

The Rasch approach generates estimates of item difficulty and person ability using a common scale. This scale is based on the logarithm of the odds of a person getting an item correct given the relative person ability and item difficulty. There is a considerable advantage in having ability and difficulty assessed on a common scale. There is also a disadvantage in that the scale that results from the modelling process uses units known as logits (logarithm of probabilities) and these units are not in common usage. However, these scales have some important properties. They are linear and do enable true comparisons across the scale. The logit scales that result from Rasch analysis have arbitrary origins based upon the mean difficulty of items. However, being linear, they enable the transformation of scores to scales that are more meaningful to participants.

An important consequence of the Rasch measurement model is that different groups of participants can be administered different versions of an instrument and, provided that there are some items in common, the two groups can receive comparable scores that take into account the differences in item difficulty.

There are many other advantages of the Rasch measurement model and it is becoming increasingly important in the assessment and reporting of educational achievement and in many other fields. A very accessible introduction to the Rasch model is given by Bond and Fox (2001).

Qualitative data

Although the study was designed primarily as a quantitative one, a considerable volume of qualitative information was gathered as a result of an evaluation in which student opinions were surveyed and their comments invited on various aspects of the new approach to the assessment of problem-solving that was being trialled. These comments were recorded in the student records system and have been reviewed and analysed.

4 Results

In this chapter the results of analyses of data collected through the Authentic Performance-based Assessment of Problem-Solving Project are presented. The characteristics of participants are summarised. Then the results of analyses of the problem-solving inventory are presented in summary form, followed by the results of more detailed analyses of the problem-solving assessment. In both cases, the outcomes of conventional analyses are shown before the findings of Rasch methods. Data that illustrate the development of problem-solving ability over multiple testing occasions are analysed as are the relationships between attitude, problem-solving ability and educational achievement. The chapter concludes with an analysis of students' evaluations of the processes of assessing key competencies.

Some of the discussion is necessarily technical. Some implications of these results are offered in chapter 5.

The results of two studies are presented in this chapter. The main study involved 43 assessments undertaken by 33 students enrolled in electronics and information technology courses at Torrens Valley Institute of TAFE. A subsequent, validation study, conducted within the Certificate IV in Assessment and Workplace Training course at Torrens Valley Institute, involved 48 participants, each of whom submitted one assessment. A further eight assessments were collected from Electronics and Information Technology Program participants, but were too late to be included in the analyses reported here. In all, 99 assessments have been completed using the problem-solving assessment instrument.

Participants

Of the 33 students who participated in the main study, 25 students submitted one piece of work for problem-solving assessment, six submitted two, and two individuals submitted three assignments, for a total of 43 problem-solving assessments. Of 25 tasks that had been identified as potential targets for problem-solving assessments, 18 were the subject of assessments.

The problem-solving assessment instrument was also tested in a validation study within the Certificate IV in Assessment and Workplace Training with a sample of 48 respondents.

Thirty of the students who participated were males and three were females. The ages of participants ranged from 17 to 50 years, with a mean of 29.45 years.

The previous educational attainment of participants varied. Three had completed Year 10, six had completed Year 11, 22 had completed Year 12, and two did not respond to this question. The post-school educational attainment also varied. Thirteen individuals had no post-school education other than the electronics and information technology course that they were undertaking, six had undertaken other VET sector courses, seven had completed other VET sector awards, three had undertaken study at university and three did not provide this information.

The distribution of work experience among participants was bimodal. Nine students reported having no work experience, six had less than a year in the workforce, two had from two to five

years' experience, and 22 people had more than five years' experience in the workforce. One student did not respond to this item. Among the participants there appeared to be a bias towards mature experienced individuals and this may have influenced aspects of the results.

The problem-solving inventory

The problem-solving inventory was not the principal focus of the Authentic Performance-based Assessment of Problem-Solving Project but was used to provide information on participants' attitudes to problem-solving. For this reason, although the results obtained from its administration have been subject to extensive analyses, only summaries of these analyses are presented here. Details of the analyses are available from the first author.

Factor analysis and classical item analysis

A factor analysis was conducted on the data gathered from the administration of the problem-solving inventory to the 33 participants from the Electronics and Information Technology Program. The analysis employed SPSS (SPSS Inc. 1995) and a three-factor model was specified. The three factors together accounted for 48.1 per cent of observed variance. This analysis clearly supported the 'personal control' (PC) subscale that is claimed by the developers of the problem-solving inventory (Heppner & Petersen 1982). However, the approach–avoidance style (AAS) and the problem-solving confidence (PSC) subscales were not clearly differentiated in the current data set. Given the modest size of this data set, it is inappropriate to revise the instrument structure, and the scales proposed by the developers have been retained for subsequent analyses. However, one item (item 32) was moved from the PSC to the PC scale. This was strongly indicated by its high loading (0.708) on the same factor as other PC items and, as was apparent on reviewing the item, the wording suggested that it was more appropriately placed on the PC scale.

Classical item analyses were conducted using SPSS. The Cronbach alpha coefficients of the scale and its three subscales were: PSI 0.72 (32 items); AAS 0.52 (16 items); PSC 0.67 (10 items); PC 0.79 (6 items). These figures suggest that the AAS and PSC subscales were not working well. Rather than explore these further with classical analysis, the instrument was reanalysed under the Rasch measurement model using Quest (Adams & Khoo 1999).

Rasch analysis

Rasch analyses were undertaken for two purposes. First, the Rasch model is able to detect deviations from expected patterns of responses for both items and respondents. Items that underfit may not be measuring the intended construct and may contaminate the measurement. Second, once a set of fitting items has been established, estimates of respondents' attitudes on the scales were generated for later comparison with problem-solving assessment scores.

The Rasch analysis indicated that six of the 32 items of the overall scale revealed poor fit and, in several stages of refinement, these items were removed from the analysis. The main basis for removing items was that they showed an Infit MS of >1.35, although item discrimination was also examined. Generally, items with high Infit MS values also showed low (<0.40) item discriminations.

A final set of 26 items was retained for further analyses. The Cronbach alpha values of the revised scale and subscales were: PSI 0.91 (26 items); AAS 0.83 (12 items); PSC 0.87 (9 items); PC 0.83 (5 items). These values indicate that the revised scale and subscales were rather more coherent than the original.

The problem-solving assessment

The problem-solving assessment instrument is a newly developed assessment tool, and requires validation. Both conventional approaches to the validation of the instrument and the use of the Rasch measurement model have been applied and the results of these analyses are reported.

The form of the problem-solving assessment instrument that was used in this study is presented in appendix 4. In the tables that follow in this section, abbreviations are used for each of the indicators in the instrument. The indicators and their abbreviations are shown in table 4.

Table 4: Major processes, indicators and abbreviations of the problem-solving assessment instrument

Major problem-solving process and indicator	Abbreviation
Representation	
Forms a correct understanding of the problem	REP01
Recognises relevant given information	REP02
Identifies the need for additional information	REP03
Recalls relevant information	REP04
Sets a realistic goal	REP05
Planning	
Plans an approach to the problem	PLN01
Recalls previous relevant or similar problem tasks	PLN02
Identifies appropriate subgoals	PLN03
Checks that required equipment is available	PLN04
Sets an appropriate time frame	PLN05
Execution	
Begins to follow the set plan	EXE01
Activates relevant knowledge	EXE02
Uses relevant skills	EXE03
Application of strategies	EXE04
Monitoring	
Checks that set plan leads toward problem goal	MON01
Response to deviations from expected progress	MON02
Reviews orginal plan	MON03
Checks problem representation	MON04
Reflection	
Reviews efficacy of problem approach	REF01
Compares current problem with previously encountered ones	REF02
Anticipates situations in which current problem approach might be useful	REF03

Factor analysis

An exploratory factor analysis was undertaken using SPSS (SPSS Inc. 1995). A scree plot suggested that three factors could be identified, although five had Eigen values >1.0. However, some caution must be exercised in interpreting factor structures in exploratory factor analysis because items that are skewed tend to have suppressed correlations with comparable items that are not skewed. This leads to the recognition of additional factors in order to account for the pattern of correlations. Certainly, there were some skewed items, and possible reasons for this are discussed later. In addition, some items had zero variances (see table 6) and these had to be removed from the factor analysis. Interpretation of the factor structure therefore is not completely clear. There appeared to be a separation between the 'execution' items and the more metacognitive ones of representation, planning, monitoring and reflection (table 5).

Table 5: Rotated factor solution for the problem-solving assessment

Item	Factor 1	Factor 2	Factor 3
REP01	.41818	.76886	–
REP02	.68305	–	–
REP03	.58550	.32448	–
REP04	.57119	.15555	.14363
PLN01	.66339	.30012	.59391
PLN02	.80128	.43937	–.26545
PLN05	.48429	.57838	–
EXE02	–.18817	.67127	–
EXE03	–.53234	.50818	.12289
EXE04	.78184	.40475	.45927
MON02	–.10229	.80097	–
MON03	.14583	.66416	–.16472
MON04	.19699	.71233	.24915
REF01	.31047	.75269	–
REF02	.22775	.67888	.24385
REF03	–.51544	.59311	–.11060

Note: Principal components extraction with Oblimin rotation. Factor loadings of <0.1 suppressed

The constraints on the variables (low variance and skewed responses), along with the lack of clear factor separation, suggest that the optimum solution may have a single factor. This proposition requires confirmatory factor analysis for which a greater number of cases are required.

Classical item analysis

The results of a classical item analysis conducted using the Scale Reliabilities command in SPSS are shown in table 6. It is worthy of note that some items had zero variance (indicated by the zero corrected item–total correlations). This is likely to reflect some bias in the candidates who volunteered for this assessment. It seems likely that more confident and more able students might have volunteered and might therefore have performed uniformly well on some of the indicators. Despite these limitations, the reliability coefficients for the scale and subscales were generally reasonably good (see table 7).

The reliability coefficients for the problem-solving assessment scale and its subscales are shown in table 7. The instrument was constructed around five major problem-solving processes, with indicators for each. A deliberate decision was taken to restrict the number of indicators to make application of this assessment tool as efficient as possible, while retaining enough indicators to provide reliable assessment. A result of having relatively few indicators for each process, and therefore subscale items, is that the Cronbach alpha values tend to be lower than would be the case for instruments with many items. An alternative way to construct the scale is to combine the representation and planning indicators into a single 'preparation' subscale, and the monitoring and reflection indicators into a single consolidation subscale. Reliability indicators for these combined subscales are also shown in table 7.

Table 6: Results of scale reliabilities

Item	Scale mean if item deleted	Scale variance if item deleted	Corrected item –total correlation	Alpha if item deleted
REP01	24.3750	19.0500	.5352	.7782
REP02	24.5625	19.7292	.2514	.7923
REP03	25.3125	19.8292	.3780	.7865
REP04	24.5625	18.9292	.4405	.7813
REP05	25.1875	21.0958	.0000	.7967
PLN01	24.1250	15.8500	.6643	.7604
PLN02	25.4375	18.9292	.5054	.7785
PLN03	25.1875	21.0958	.0000	.7967
PLN04	25.1875	21.0958	.0000	.7967
PLN05	25.5625	19.0625	.4085	.7832
EXE01	25.1875	21.0958	.0000	.7967
EXE02	24.3125	20.3625	.2000	.7935
EXE03	24.3125	20.7625	.0696	.7985
EXE04	25.1875	14.4292	.7812	.7465
MON01	25.1875	21.0958	.0000	.7967
MON02	23.5000	19.7333	.1869	.7984
MON03	25.1875	18.5625	.3178	.7920
MON04	25.6875	18.3625	.5574	.7739
REF01	24.8125	18.4292	.5629	.7739
REF02	25.6250	17.4500	.5128	.7751
REF03	25.2500	21.0000	.0145	.7987

Table 7: Scale reliability indices for the PSA and subscales

Scale	Cronbach alpha	N items
PSA	0.795	21
Representation	0.655	5
Planning	0.450	5
Execution	0.263	4
Monitoring	0.605	4
Reflection	0.506	3
Preparation	0.725	10
Consolidation	0.616	7

Note: PSA=problem-solving assessment scale

The results of the classical item analysis indicate that the problem-solving assessment scale had satisfactory measurement properties despite the lack of variance of some items. However, more detail is available from Rasch analyses.

Rasch analysis

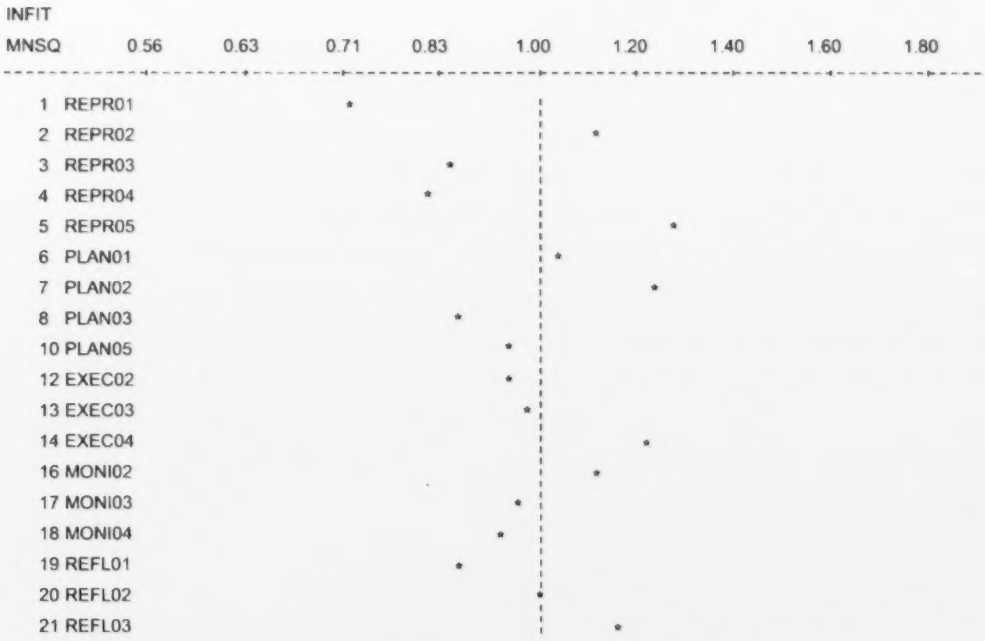
Rasch analyses were undertaken on the problem-solving assessment scale using both Quest (Adams & Khoo 1999) and RUMM (Sheridan et al. 1997). The purposes of these analyses were to evaluate the measurement properties of the scale, to examine the performance of individual items in the scale and to generate psychometrically meaningful estimates of students' abilities.

Scale coherence and item fit

Initial analyses were conducted using Quest. These showed that the instrument as developed had quite good scale properties. The item reliability estimate, a measure of the extent to which items were differentiated by participants of different ability, was 0.85. The person separation index, a measure of the extent to which the items of the scale were able to discriminate among participants of different abilities, was 0.81, approximately equivalent to the Cronbach alpha (0.795).

Item fit was assessed in Quest. Two items showed a significant level of misfit. Items that fit very well have Infit MS values of 1.0, with acceptable values ranging from 0.71 to 1.30. Item 9, the fourth planning indicator, had an Infit MS value of 1.98 indicating that this variable was influenced by a factor that was not reflected in other items. Item 15 had a very low Infit MS value (0.28), indicating that this item added no information to the scale that was not already provided by other items. These items were removed and the analysis repeated. A diagram showing the fit parameters of the remaining items is presented in figure 1. For the reduced scale, the item reliability estimate was 0.84 and the person separation index was 0.81.

Figure 1: Fit parameters (Infit MS) of the problem-solving assessment scale



Additional analyses were conducted using RUMM. This was selected for two reasons. First, while Quest uses a marginal maximum likelihood method, RUMM uses a paired comparison algorithm to calculate estimates. This is more time consuming, but in data sets of modest sizes, it produces more precise estimates. Second, RUMM produces superior graphical displays of certain outputs that illustrate better the characteristics of items.

Item locations and thresholds

The locations of items, an indication of their average difficulty, were estimated, as were item thresholds. In items with multiple responses, the ability level at which a person is likely to move

from one response level to the next (for example from a 1 to a 2) is the threshold between those response levels. Dichotomous items have a single threshold (equal to the item location, or difficulty), while items with four response options (the maximum number used in the problem-solving assessment instrument) have three thresholds. The range of item locations and thresholds provides an indication of the range of ability levels that the instrument is able to measure with reasonable precision. Item locations and thresholds for the problem-solving assessment scale are shown in table 8. Locations varied from –3.1 to +3.3, a range of 6.4 logits. Thresholds varied from –5.1 to +4.2, a range of 9.3 logits. These ranges would permit the reliable measurement of problem-solving ability over a very broad span. With instruments such as the problem-solving assessment instrument, one might expect variations in person ability from approximately –3 to approximately +3, with only a small proportion of people outside this range.

Table 8: Estimates of PSA item locations and thresholds

	Location		Thresholds		
Statement	Estm	SE	1	2	3
REPR01	–0.852	0.51	–1.660	–0.044	–
REPR02	–1.778	0.41	–4.704	1.149	–
REPR03	–0.805	0.75	–0.805	–	–
REPR04	–1.640	0.38	–4.927	1.647	–
PLAN01	–0.076	0.27	–5.074	1.839	3.007
PLAN02	0.184	0.52	0.184	–	–
PLAN03	0.013	0.56	0.013	–	–
PLAN04	–2.037	1.10	–2.037	–	–
PLAN05	1.197	0.42	1.197	–	–
EXEC02	0.899	0.29	1.406	0.391	–
EXEC03	–2.752	0.65	–4.988	–0.516	–
EXEC04	2.273	0.24	2.872	1.673	–
MONI01	–3.107	1.52	–3.107	–	–
MONI02	1.099	0.23	0.060	2.063	1.176
MONI03	2.328	0.32	1.070	3.585	–
MONI04	1.380	0.44	1.380	–	–
REFL01	0.042	0.38	–2.733	2.817	–
REFL02	3.308	0.32	2.372	4.245	–
REFL03	0.324	0.54	0.324	–	–

Note: PSA=problem-solving assessment scale

The probability of a person of a given ability being graded in a particular response category on an indicator is illustrated by considering a category probability curve for item 1, 'Forms a correct understanding of the problem'. This is shown in figure 2. People whose overall problem-solving ability lay below about –1.7 were most likely to be graded as category 0, 'Misunderstands the problem' on this indicator (see the curve labelled '0' in figure 2). People with somewhat higher ability, between –1.7 and 0, were likely to be graded as category 1, 'Forms a partial understanding using some given information'. People of higher ability, above 0 logits on the scale, were likely to be graded as category 2, 'Forms a complete understanding using all relevant factors'.

Not all items showed the expected pattern of response probabilities. For example, item 12, 'Activates relevant knowledge' revealed the pattern shown in figure 3. In this case, people with a problem-solving ability below +1.0 logits were likely to have been graded as category 0, 'Does not recall relevant knowledge or recalls incorrect information'. Above this level (+1.0 logits), the most likely performance level was category 2, 'Activates and uses relevant knowledge'. The middle category (category 1, 'Activates some, but not all, relevant knowledge') was less likely than the other two to be used at any ability level. This could lead to the interpretation that the item was

effectively behaving as a dichotomous one and that it would be possible to simplify the indicator by removing the middle performance level. This course of action is not recommended in this case, and possible modifications to the instrument are discussed in chapter 6, Conclusions and future directions.

Figure 2: Category probability curve for item 1

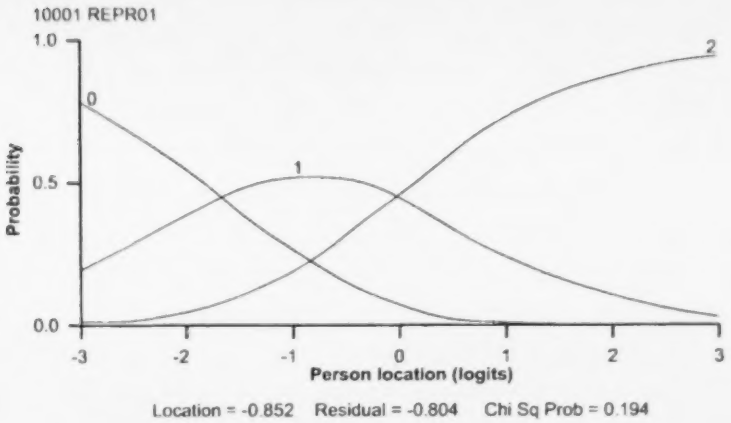
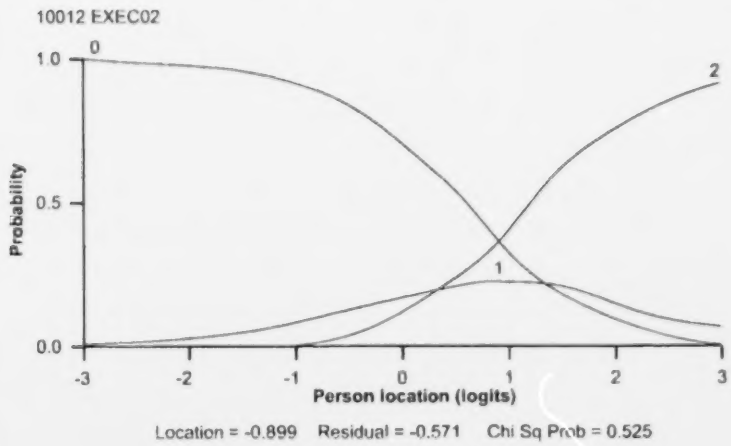


Figure 3: Category probability curve for item 12



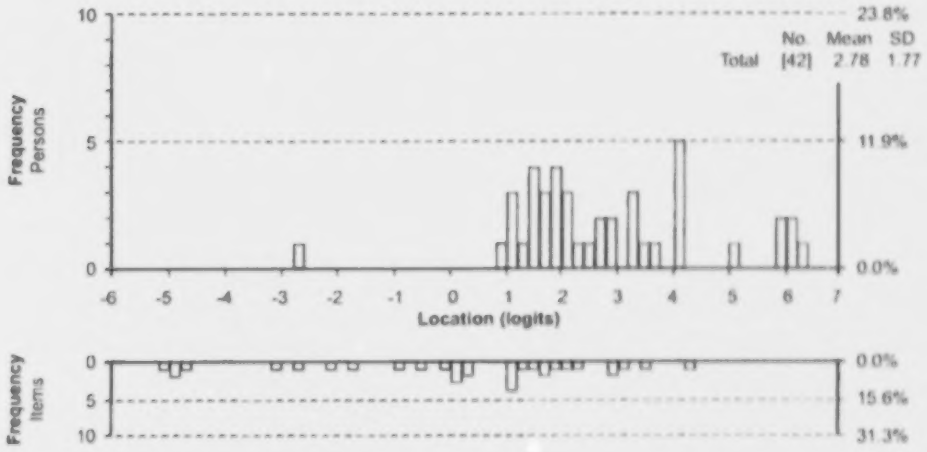
Person abilities

An issue of concern in measurement is the match between the items that are used to assess a trait, in this case problem-solving ability, and the distribution of that trait in the target group. The match for the test sample is illustrated in figure 4. The top plot is the distribution of person ability and the lower plot is the distribution of item thresholds.

The zero position on the scale is the mean difficulty of items. In most cases person ability lay above this mean. This may indicate that most items were too easy for the target population or it may reflect a tendency for more able students to volunteer for this assessment. The fact that there

was one case near –3 logits suggests that the scale may be reasonably targeted, although the several cases with scores in excess of the highest item threshold suggest that there is a need either for additional indicators at the high end of the scale or for more rigorous performance criteria to be applied to existing indicators. The instrument does need to be tested with a wider range of candidates before a conclusion on its targeting can be reached.

Figure 4: Distributions of person ability (top) and item thresholds for the PSA



Note: Grouping is set to an interval length of 0.20, making 65 groups
PSA=problem-solving assessment scale

The ability distribution among participants is of interest in establishing the number of performance levels that can be supported. The single very low scoring case may be an outlier, but it may reflect the likely performances of individuals with low problem-solving ability. Among the distribution of the more successful participants, there appear to be three identifiable nodes, one centred around 1.7 logits, one around 3.5 logits and another at around 5.5 logits. This suggestion is tentative and its verification would require many more cases.

Reporting student problem-solving ability

The abilities shown in the above data are expressed in the natural unit of Rasch measurement—the logit. Although in principle the scale is open ended, it is relatively compressed, with values typically ranging from –3 to +3, and often expressed to two decimal places. These numbers are generally not meaningful to non-specialists, and a metric for this ability is required in units that may be more attractive to practitioners, students and employers. Because it is an interval scale, it can be transformed linearly. A scale must be found that will communicate ability clearly and simply and not lead to misunderstanding. This suggests that measurements on the scale should be expressed as integers. However, if values were expressed in units for which common values were below 100, it might be interpreted as a percentage scale.

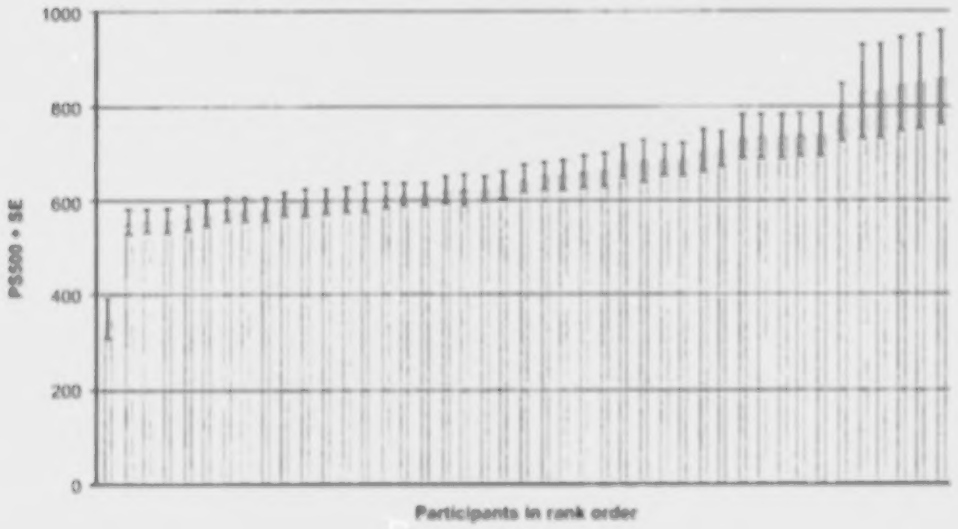
In many other scales, it has become common practice to transform the measures found from Rasch analyses into scales with means of 500 and standard deviations of 100. This practice is suggested for the problem-solving assessment scale and has been implemented. The transformed scale is labelled the problem-solving 500 (PS500) scale. Problem-solving scores and their standard errors in both the natural Rasch unit (logits) and on the transformed (PS500) integer scale are shown in

BEST COPY AVAILABLE

table 9. These estimates were obtained using RUMM and checked against the estimates found using Quest. The Quest estimates were similar, but in most cases the standard errors using RUMM were approximately 20 per cent lower than those resulting from the Quest analysis. In larger samples, this difference is likely to be less pronounced.

An advantage of using the Rasch measurement model is that it has produced estimates of known precision—a precision indicated by the standard errors of the estimated problem solving ability. Figure 5 depicts the PS500 data from table 9, sorted in ascending order, and showing the standard errors of the estimates. For estimates near the scale mean, standard errors are at their smallest and the precision is at its highest, but for abilities near the extremities of the distribution, the precision is lower.

Figure 5: Participants' problem-solving ability for all assessments on the PS500 scale



The PS500 scale may be a useful metric for reporting estimates of problem-solving ability. The form in which such estimates could be reported is discussed in chapter 5.

Table 9: Student problem-solving ability (logits and transformed values)

		Logit values		PS500 values	
Participant ID	Occasions	Ability	SE	Ability	SE*
STE001	1	6.074	–	847	99
STE002	1	2.564	0.562	646	32
STE003	1	1.824	0.549	604	31
STE003	2	2.142	0.471	622	27
STE004	1	3.273	0.593	687	34
STE005	1	4.158	0.796	738	46
STE006	1	1.799	0.444	603	25
STE006	2	2.712	0.537	655	31
STE007	1	0.958	0.456	550	26
STE007	2	1.413	0.438	581	25
STE007	3	2.899	0.570	666	33
STE008	1	-2.641	0.741	349	42
STE009	1	1.413	0.438	581	25
STE010	1	4.118	0.810	735	46
STE010	2	3.273	0.593	687	34
STE011	1	1.689	0.470	596	27
STE012	1	2.639	0.575	662	33
STE013	1	3.226	0.588	684	34
STE013	2	6.304	–	850	99
STE014	1	3.249	0.776	686	44
STE014	2	6.174	–	853	99
STE015	1	3.607	0.799	708	46
STE016	1	5.022	1.087	787	61
STE016	2	4.194	0.792	740	45
STE017	1	5.824	–	833	99
STE018	1	2.166	0.556	624	32
STE019	1	5.824	–	833	99
STE020	1	2.000	0.452	614	26
STE021	1	1.413	0.438	581	25
STE022	1	3.668	0.668	710	38
STE023	1	1.247	0.466	571	27
STE024	1	2.167	0.485	625	28
STE025	1	1.605	0.439	592	25
STE026	1	2.000	0.452	614	26
STE027	1	1.121	0.477	564	27
STE028	1	4.146	0.797	737	46
STE029	1	2.340	0.519	634	30
STE030	1	4.194	0.792	740	45
STE031	1	1.722	0.472	598	27
STE032	1	1.026	0.443	559	25
STE032	2	1.026	0.443	559	25
STE032	3	1.973	0.469	613	27
STE033	1	2.677	0.507	653	29

Note: * SE values of 99 are extrapolated from SEs of other scale estimates.

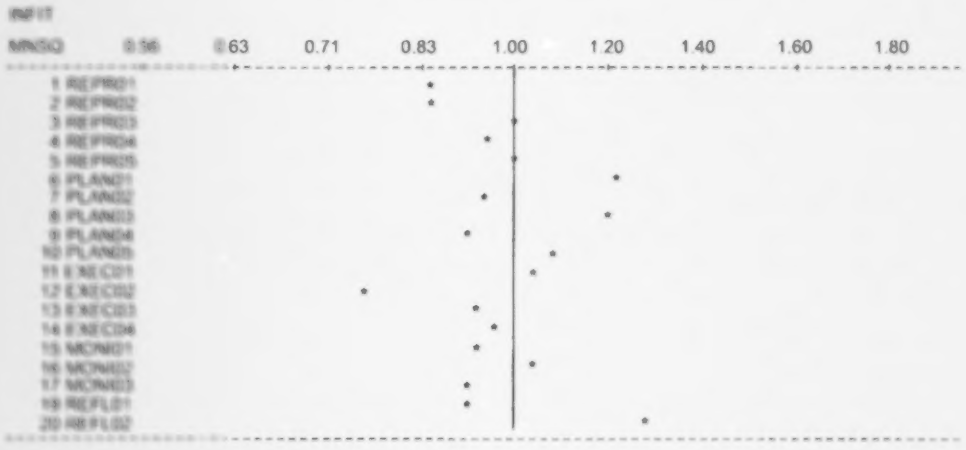
The validation study

In order to begin to test the applicability of the problem-solving assessment instrument in a broader range of courses, a validation study was undertaken. This involved 48 individuals who were enrolled in the Certificate IV in Assessment and Workplace Training. They were all doing the module 'Plan and promote a training program'. Each participant completed one assessment using the problem-solving assessment instrument and the results of this trial were recorded and analysed separately from those of the main data set. Here summaries of the Rasch analyses are presented.

Scale coherence and item fit

The Cronbach alpha for the scale found using the 48 cases in the validation sample was 0.77. Quest (Adams & Khoo 1999) produces two overall indicators of scale coherence. The item reliability estimate was 0.82 and the person separation index was 0.75. The corresponding values in the main sample were 0.85 and 0.82 respectively. In the validation sample, item parameters could not be estimated for items 18 and 21 (MONI04 and REFL03). The remaining indicators were found to fit a single coherent scale. In the main sample, two indicators (9 and 11) were found not to fit. The fit parameters for the remaining items in the validation sample are shown in figure 6.

Figure 6: Fit parameters (Infit MS) of the problem-solving assessment in the validation study



Item locations and thresholds

Item locations and thresholds were estimated using the validation sample. For the scale to have good measurement properties, not only must overall scale indices remain consistent when tested with a new sample, but the items should also have consistent locations. Item locations were estimated for the original Electronics and Information Technology Program sample of 43 participants, for the validation sample of 48 cases, and for a combined sample. The combined sample had 99 cases, as eight new Electronics and Information Technology Program cases became available for analysis in time for comparison with the validation sample. The locations and standard errors of these estimates are shown in table 10.

Two features of this table are of interest. First, standard errors of the item locations varied. As expected, the smaller samples had somewhat higher standard errors for given indicators. For many indicators, the standard errors of locations were between 0.15 and 0.35 logits for the larger data set, but some had rather higher standard errors (> 0.5), with one of the planning indicators (PLAN04) having an error term of 0.74. High standard errors are associated with low measurement

precision, and using these items to compute scores for individuals would lead to reduced precision in individuals' scores. These relatively high standard errors are an indication that, although the item fits the scale, it was not used consistently by raters, and this suggests the need for close scrutiny of the indicator and its performance criteria.

The second notable feature of the indicator location data is the variation in item locations across data sets. Ideally, these indicators would be equal across samples. Some tolerance must be allowed for sampling variations, but several items (REPR02, REPR04, EXEC03, MONI03 and REFL02) had location variations greater than would be expected on the basis of the standard errors for the estimates in each sample. This suggests a systematic difference in the way different sample groups interpreted and responded to indicator performance criteria.

Table 10: Estimates of item locations for all data, E&IT data and validation data

	All data (N=99)		E&IT data (N=43)		Validation data (N=48)	
Indicator	Location	SE	Location	SE	Location	SE
REPR01	−.555	0.255	−0.852	0.51	−.627	0.347
REPR02	−.916	0.232	−1.778	0.41	−.218	0.352
REPR03	−1.117	0.572	−0.805	0.75	−1.312	1.010
REPR04	.372	0.228	−1.640	0.38	1.603	0.348
REPR05	−1.177	0.569	–	–	−.043	0.588
PLAN01	.112	0.158	−0.076	0.27	.529	0.227
PLAN02	−.189	0.396	0.184	0.52	−.592	0.752
PLAN03	.276	0.338	0.013	0.56	.721	0.459
PLAN04	−1.794	0.740	−2.037	1.10	−1.429	1.085
PLAN05	.621	0.306	1.197	0.42	.133	0.552
EXEC01	−.632	0.470	–	–	.562	0.480
EXEC02	.477	0.220	0.899	0.29	−.558	0.334
EXEC03	−1.414	0.260	−2.752	0.65	−.659	0.340
EXEC04	2.126	0.160	2.273	0.24	1.847	0.268
MONI01	−.761	0.489	−3.107	1.52	.232	0.534
MONI02	.920	0.153	1.099	0.23	.390	0.225
MONI03	1.489	0.194	2.328	0.32	673	0.326
MONI04	.219	0.362	1.380	0.44	–	–
REFL01	.564	0.229	0.042	0.38	−.598	0.347
REFL02	2.070	0.182	3.308	0.32	−.654	0.338
REFL03	−.691	0.502	0.324	0.54	–	–

Note: Blank cells represent indicators for which locations could not be estimated for that sample
E&IT = electronics and information technology sample

The development of problem-solving ability

The problem-solving assessment instrument is a tool that specifically and overtly specifies core problem-solving processes, indicators and performance levels. The processes used in the administration of this instrument involved student self-assessment followed by facilitator validation and discussion. Given these conditions, it was hypothesised that students would develop explicit knowledge about problem-solving and that their performance would be enhanced as a direct result of the assessment process. In short, it was hypothesised that the assessment regime would lead directly to improved performance.

The design of the study anticipated that many students would participate and that most would undertake assessments on three separate occasions. Under these conditions, the use of a hierarchical modelling of performance using Hierarchical Linear Modeling (HLM) (Bryk, Raudenbush

& Congdon 1996) would have enabled the hypothesis of enhanced performance to be tested robustly. However, due to the voluntary nature of participation and the relatively compressed timeframe of the project, few students undertook problem-solving assessments on multiple occasions. Indeed, only six students undertook assessments on two occasions and a further two students undertook assessments on three occasions. These numbers are inadequate to justify the level of statistical analysis that was envisaged as meaningful statistical significance could not be established.

The change in student performance is depicted in two figures. Figure 7 shows the change in problem-solving performance of the six students who undertook problem-solving assessments on two occasions. Four of the students showed improvements in performance, but two showed slight, although non-significant, declines.

Figure 7: Change in problem-solving performance over two assessment occasions

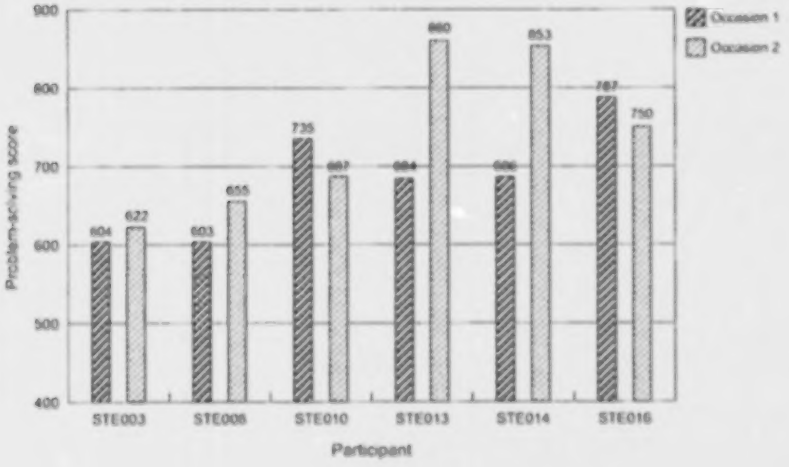
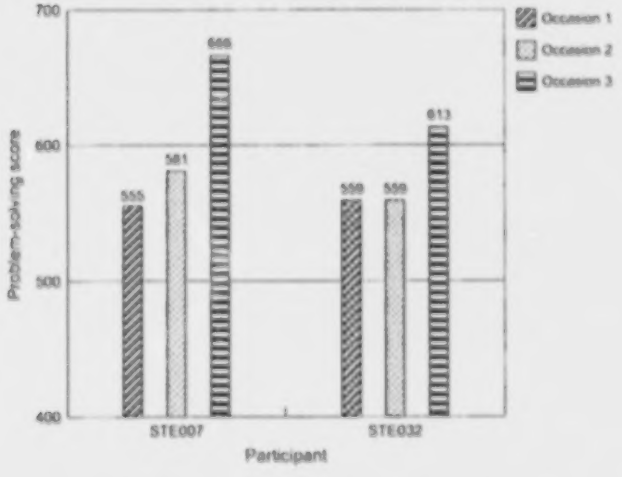


Figure 8 shows the change in problem-solving performance for the two participants who completed three problem-solving assessments. In both cases, there appears to be a trend of increasing performance, but because of the small sample size, no statistical support can be provided for this claim.

Figure 8: Change in problem-solving performance over three assessment occasions



Problem-solving and educational achievement

One of the hypotheses that informed this research was that high levels of problem-solving performance would be associated with enhanced educational achievement. In addition, extensive research on affective dimensions of performance has shown that learners who have positive perceptions of their abilities (high perceived self-efficacy) are likely to perform better than those with lower views of their abilities (Bandura 1989; Dweck 1986). Thus, both traditionally cognitive factors, such as knowledge use and problem-solving processes, and attitudinal factors are implicated in models of student performance. In much research on educational achievement, prior achievement is a very strong predictor.

In the Authentic Performance-based Assessment of Problem-Solving Project, data on attitude to problem-solving were obtained through the administration of the problem-solving inventory. Data on problem-solving ability were obtained from the problem-solving assessment, and students' grades on their current modules provided information on educational achievement. Data on prior achievement were not gathered.

In order to test the proposition that educational achievement was influenced favourably by positive attitudes and by problem-solving ability, it was desirable to construct a path model and to compare its predicted relationships with the data that emerged from the study.

In vocational education and training in Australia, assessment is competency based, and in general learners are graded dichotomously, either as having achieved the specified level of competence or as not having achieved it. In the Electronics and Information Technology Program at Torrens Valley Institute, a 'graded competencies' approach is used. In this method, a minimum competence level is specified for all assessed tasks in a module and students who achieve this level are awarded a pass in that module. However, students may seek to gain higher grades. Additional, credit and distinction, objectives are specified in each module. These objectives relate to the abilities of learners to demonstrate an integrated or holistic understanding and application of core module competencies. Thus, they are not additional and separate competency objectives (that is, more but different), but rather they are an attempt to assess enhanced quality of learning.

In conventional approaches to assessment all students are assessed against established criteria and may be graded according to those criteria or be scaled using a norm-referenced approach. These approaches lead to the identification of many levels of achievement. Most implementations of competency-based assessment lead to two performance levels: competent and not yet competent. Under the graded competency approach used in the Electronics and Information Technology Program, students may elect to have additional competencies assessed, and four levels are available: fail (not yet competent), pass, credit and distinction. The higher grades reflect both an intention to be assessed and success on that additional assessment. Achievement under this approach is likely to be related to indications of ability and, perhaps more strongly than in conventional assessment models, to attitudinal factors.

In order to investigate the relationships between attitude, problem-solving ability and educational achievement, a correlation analysis was undertaken. The variables used in this analysis included:

- the three subscales of the problem-solving inventory: approach–avoidance style (AAS), problem-solving confidence (PSC) and personal control (PC)
- the overall problem-solving inventory attitude score (PSI)
- the problem-solving ability score, included in the analysis for students who had completed more than one problem-solving assessment task (PSA)
- educational achievement, indicated by the grade a student had achieved in the module as a whole (Grade). A fail was coded as 0, a pass coded as 1, a credit coded as 2 and a distinction

coded as 3. Some students had not completed their modules at the time of data analysis, and their grades were coded as missing.

The results of this analysis are displayed in table 11.

Expected patterns of correlations were observed among the three attitude scales. However, negative, although non-significant, correlations were observed between the attitude scales and both problem-solving ability and grade. An expected positive and significant correlation was present between problem-solving ability and grade. The value of 0.46 for this correlation was high considering that the grade variable had a truncated range, since only three levels of this variable are present in the data. The unexpected relationships between attitude and both problem-solving ability and educational achievement suggest the need for alternative and more detailed analyses.

Table 11: Correlations between attitude, problem-solving ability and educational achievement

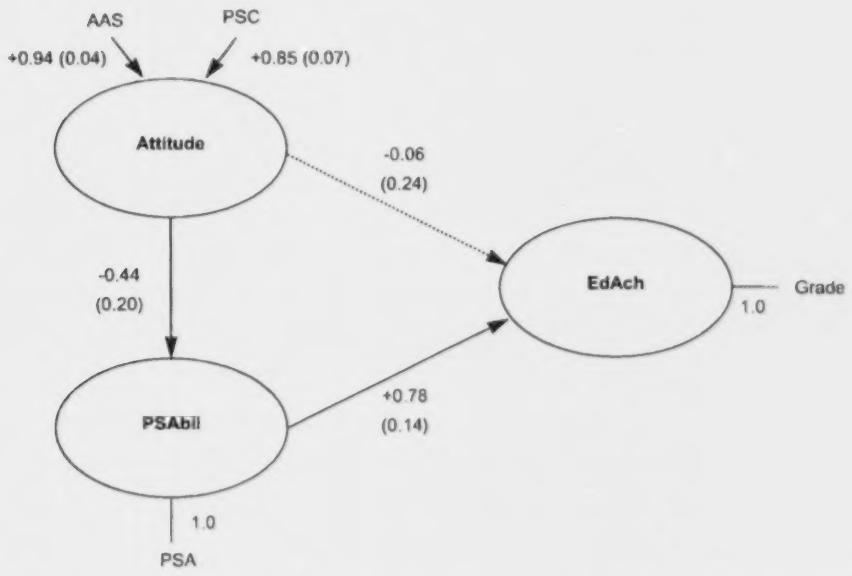
	AAS	PSC	PC	PSI	PSA	Grade
AAS	1.0000					
	33					
PSC	**0.7164**	1.0000				
	33	33				
	0.0000					
PC	0.3357	**0.5071**	1.0000			
	33	33	33			
	0.0560	0.0030				
PSI	**0.8254**	**0.8981**	**0.5768**	1.0000		
	33	33	33	33		
	0.0000	0.0000	0.0000			
PSA	–0.2166	–0.0682	–0.0534	–0.0810	1.0000	
	33	33	33	33	33	
	0.2260	0.7060	0.7680	0.6540		
Grade	–0.1973	–0.2088	0.1888	–0.1863	**0.4574**	1.0000
	26	26	26	26	26	26
	0.3340	0.3060	0.3560	0.3620	0.0190	

Note: Each cell has the correlation coefficient, the number of cases and the level of significance of the coefficient. Where $p<0.05$, the coefficients are shown in bold

In order to investigate these relationships further a latent variable path analysis was conducted using PLS Path (Sellin 1989). The path model was constructed around three latent constructs: attitude, problem-solving ability (PSAbil), and educational achievement (EdAch). The manifest variables were those used in the correlation analysis. The path model is shown in figure 9.

In the model, the personal control scale did not load onto the attitude latent construct, and so was deleted from the model. The path from attitude to educational achievement (EdAch) was not significant. Surprisingly, the path from attitude to problem-solving ability (PSAbil) was negative. There was a strong positive path from problem-solving ability to educational achievement.

Figure 9: Path model showing the relationships between attitude, problem-solving ability and educational achievement



Note: Latent constructs are shown in ellipses and manifest variables in rectangles. Standardised path coefficients are shown, with jack-knife standard errors in parentheses

Evaluation by students

In order to assess the acceptability of the approach to problem-solving ability assessment and the instrument used for this purpose, student opinion was canvassed through a survey instrument and their anonymous comments were sought via a feedback form or email. Their responses to the survey are summarised and their comments analysed to identify issues that are important to students.

Responses to the student evaluation survey

Of the 29 students who responded to the survey, 22 had completed key competencies assessments and seven had not. The evaluations of both groups are reported below.

Evaluation by students who had completed key competencies assessments

One of the issues of interest to the researchers was the extent to which both participants and non-participants felt informed about the purposes and processes of assessments of key competencies. Table 12 shows the responses of students who had undertaken key competencies assessments to questions concerning the adequacy of information about the process and their intentions to participate in future assessments. There is a strong indication that these students did feel well-informed and, among existing participants, an intention to seek further assessments of their key competencies was apparent.

Table 12: Perceptions of the adequacy of information about key competencies assessments (participants)

Item	Yes	Largely	Partly	No
Do you feel well-informed about the key competencies assessment process?	10	8	2	0
Is there enough information available to inform people about key competencies assessment?	12	6	3	0
Do you think you will apply for more key competencies assessments in the future?	13	5	3	0

The data presented in table 13 show that students who had undertaken key competencies assessments generally strongly endorsed employment outcomes, skills recognition and personal development as reasons for seeking such assessments. Some lecturers observed that being asked by a lecturer to undertake a key competencies assessment did not feature as a reason.

Table 13: Reasons for seeking key competencies assessments (participants)

	Yes	Largely	Partly	No
To get these skills formally recognised	18	0	0	1
To help give you an edge when applying for a job	16	0	0	3
To help you prove and explain your skills at a job interview	14	0	0	3
To get recognition for the extra skills you use in this flexible learning program	14	0	0	3
To better understand your skills in this area	13	0	0	4
To improve your skills in this area	15	0	0	3

The items shown in table 14 further explored participant students' perceptions of the use and administration of the key competencies assessments. Support from facilitators was regarded very favourably, but two issues that received less favourable endorsement were the clarity of the process and the extent to which participation enabled students to understand and improve their skills. Despite these reservations, most students indicated that they would recommend the process to others.

Table 14: Other perceptions of key competencies assessments (participants)

	Yes	Largely	Partly	No
Have you received adequate assistance from facilitators?	17	2	2	0
Is the assessment process clear and easy to follow?	7	9	5	1
Have these assessments helped you understand and improve your skills?	10	6	5	1
Would you recommend it to other students?	15	5	2	0

Table 15 shows that students who had participated in key competencies assessments found the revised process using the problem-solving assessment instrument informative, but clearly experienced some challenges in having their key competencies assessed using this tool. There was general agreement that using it was superior to using previous assessment forms and participants felt that it would help them in demonstrating and explaining their skills.

Table 15: Perceptions of the key competencies assessment process using the problem-solving assessment instrument

	Yes	Largely	Partly	No
Do you think it is:				
... informative?	10	5	0	0
... easy to use?	5	9	1	1
... effective in helping you understand your skills?	7	6	3	0
... better than the other assessment forms?	4	3	1	0
Do you think key competencies assessments will help you prove and explain your skills at a job interview?	14	5	3	0

Evaluation by students who had not completed key competencies assessments

The views of students who had not taken part in key competencies assessments were also sought, since this is a group whose needs must be addressed if greater levels of participation are to be achieved in a regime of voluntary participation. Table 16 shows that these students were as well-informed about the process as those who did participate. The key difference was that they elected not to participate, but the reasons for this were not reflected in their responses to the survey. Some indications can be identified in the comments the students made, and they are analysed in the next section.

Table 16: Perceptions of the adequacy of information about key competencies assessments (non-participants)

	Yes	Largely	Partly	No
Were you given enough information to help you make an 'informed decision' about doing it or not?	5	1	1	0
Have you given it much thought?	0	2	5	0
Do you know how to get more information or who to ask about key competencies assessment?	6	0	1	0
Do you think you might consider key competencies assessment in the future?	5	1	1	0

Analysis of student comments

Students were invited to comment anonymously via a feedback form or email. For this reason, it was not possible to associate comments with the IDs of students who had undertaken a problem-solving assessment. In fact, 20 students responded to the request for comments.

The full text of comments received is reproduced in appendix 5. For the current analysis, all comments were read, issues raised by students were highlighted and they were classified using headings that emerged from the comments. The classification is presented in table 17.

Table 17: Classification of comments made by students

Issue raised	Sample comment
Outcomes: personal	Doing these assessments have partly helped me understand and improve my skills.
Outcomes: employment	... make future job employers aware of your skills ...
Outcomes: recognition	I think that key competencies is a good way of being recognised for things that you do but are not necessarily recognised in any other way.
Performance levels	I find it difficult to differentiate between some of the levels.
Process	It was difficult to know when a key competency could be attempted.
Problem-solving assessment instrument	The new problem-solving trial has helped much more in that it breaks down each section of the process and allows for detailed discussion each time.
Terminology	Some of the wording in the key competency checklists are difficult to understand and it is hard to know what exactly they require.
Time	I don't think that I'll be doing any more because it takes too much time.

The issues raised most often by students were about outcomes expected as a result of having key competencies assessed. Of the expected outcomes, personal development and recognition were the most common. Comments on the assessment process were also common. Concerns about the process related to a lack of clarity about when and for which assignment tasks it would be appropriate to undertake a key competency assessment. The problem with levels of performance was identified by several students as a cause of confusion. Most of these comments appear to relate to past key competencies assessment practices, but some were also directed at the performance levels of the indicators of the problem-solving assessment instrument. A final critical comment related to the terminology used in the instrument. Some students found this difficult to understand.

Thus, the comments made by students suggest that they were well-informed about key competencies assessment, and they were well-supported and well-aware of the benefits of having key competencies recognised. However, they were not as clear about the process and the terminology used, and the criteria differentiating levels of performance remained a difficulty.

BEST COPY AVAILABLE

5 Discussion of results

The somewhat technical results of this study are presented in chapter 4. Here, the implications of those results are outlined.

Participants

The key question in relation to those people who volunteered to participate in this study relates to the generality of any conclusions that can be drawn from the Authentic Performance-based Assessment of Problem-Solving Project.

First, the study was conducted in the Electronics and Information Technology Program at Torrens Valley Institute of TAFE. The program has a history of innovation in course delivery and implements a flexible, student-centred delivery strategy. Key competencies have been an element of course delivery since the program's inception and they have been assessed specifically since 1998. This raises the question:

- Would similar results be achieved in similar courses offered by other providers who have pursued more conventional course delivery strategies?

The Electronics and Information Technology Program is a technical one, and many of its graduates can be expected to work in fault-finding and repair jobs. The course modules may therefore place greater emphasis on problem-solving than do other courses. This leads to the question:

- Would similar results be achieved in other courses, either technical ones in different areas, for example automotive, or non-technical courses, for example human services courses?

The students who took part in the study reflected a range of ages and backgrounds. However, it is possible that a greater proportion of more able students submitted work for problem-solving assessment. The age and experience distributions of participants were bimodal. Some participants were relatively young recent school-leavers with limited work experience, while others were somewhat older people with considerable experience in the workforce. There was no significant relationship between age and problem-solving performance ($r=0.071$, $p=0.704$). However, the possible bias towards high-ability participants raises the question:

- Would similar results be achieved for students of all ability levels?

The issue of possible ability bias in the sample of participants has implications beyond the generality of the results of the study. If there is no ability bias, the problem-solving assessment instrument would appear to be targeted at too low a level of problem-solving ability. Conversely, if there is an ability bias, the instrument may be adequately targeted. The possible existence of an ability bias must therefore be established before a conclusion can be reached about the targeting of the instrument, and therefore before any decisions are made about revising it.

The problem-solving inventory

The problem-solving inventory was used to assist in developing an understanding about the role of attitudes to problem-solving in problem-solving performance. The instrument was reported to have

reasonable psychometric properties, although the methods used to establish this, exploratory factor analysis and classical item analysis, are not able to provide the penetrating analyses available using item response theory methods.

As it was administered in the Authentic Performance-based Assessment of Problem-Solving Project, the problem-solving inventory did not show the same degree of separation of its three subscales—approach–avoidance style, problem-solving confidence, and personal control—as originally reported. However, this administration used a modified form of the instrument and was trialled on only 33 participants. For these reasons, some caution is warranted in interpreting its structure. The personal control factor was apparent, but the other two seemed to form an undifferentiated factor.

The problem-solving assessment tool

The problem-solving assessment tool was the central focus of the Authentic Performance-based Assessment of Problem-Solving Project, and was developed to test the hypothesis that an instrument designed to assess individuals' application of a set of generally accepted key problem-solving processes would provide both a valid and a reliable basis for measuring problem-solving ability.

The tool was developed following a wide-ranging review of theories of problem-solving and is thought to reflect high levels of both construct and content validity.

The test sample

Conventional analyses, using exploratory factor analysis and classical item analysis, suggested that the scale formed by the items of the problem-solving assessment instrument is a coherent measure of problem-solving ability. However, there are some caveats. Some dichotomous items showed no variance and this is thought to reflect a generally high ability sample of participants. However, it may be necessary to revisit these items and to provide additional performance levels for them.

Analyses based on the Rasch measurement model also provided good support for the instrument. Reliability indices for both items and persons, of approximately 0.80 to 0.85, suggest that the instrument is robust. Analyses of response patterns to individual items suggest that most items are quite sound. Some departures from expected category response patterns were found. These may reflect the modest sample size of this study, but it may be necessary to revise the number of performance levels for some items and the descriptions of those performance levels. Such revisions, however, should await trials of the instrument with larger and more diverse samples of respondents.

The validation sample

The overall problem-solving assessment scale indicators found for the validation study sample were quite similar to those computed from the original electronics and information technology test sample. This is encouraging and does suggest scale coherence. However, variations in the precision of item locations within scales suggest inconsistency in the interpretation of at least one indicator. There were also variations in indicator locations between the test and validation samples. These indicate that there were systematic differences between the two groups in interpretations of the performance criteria of the indicators. This is an undesirable feature of the instrument and suggests the need to revise the performance criteria specified for each indicator to improve the consistency of their application.

The several analytical approaches taken in examining the properties of the instrument have provided a wealth of information than can assist in the revision of indicators.

The development of problem-solving ability

The problem-solving assessment instrument appears to be a robust tool for the assessment of problem-solving ability. It was implemented in a two-stage assessment process in the Authentic Performance-based Assessment of Problem-Solving Project. In the first stage, students were asked to use the instrument (and were given detailed descriptions of the performance criteria on all indicators to assist them in its use), to self-assess their ability and to identify evidence to support that assessment. In the second stage students presented their work, along with their self-assessment, to a lecturer who made a judgement about their performance on the basis of the evidence the students were able to present.

While it is possible to use the problem-solving assessment instrument as a summative assessment device, and some practitioners may choose to take this approach, the method adopted in the project was a deliberate attempt to inform students about problem-solving. Through the use of the instrument, students were expected to develop explicit knowledge of the major processes that constitute problem-solving and, by locating evidence for each of the indicators of these processes, to incorporate that general knowledge of problem-solving into their own experience.

The trends, illustrated in figures 7 and 8 in chapter 4, suggest that students' knowledge of problem-solving was enhanced through this approach. Further support for this proposition is found in the comments made by students. Several participants commented that, although they believed that they had good problem-solving skills before commencing their courses, this assessment approach enabled them to become more knowledgeable about problem-solving and to talk about it with greater assurance. Despite these favourable assessments by most students, some commented upon a lack of clarity about the process, and some work remains to be done to further integrate key competencies assessment processes into activities for routine assessment of units of competency. In addition, some students were confused about performance levels and about some of the terminology used to describe problem-solving indicators.

Problem-solving and educational achievement

It was hypothesised that enhanced problem-solving ability would be related to high levels of educational achievement. It was also hypothesised that favourable attitudes to problem-solving would be associated both with problem-solving ability and, in particular, with educational achievement.

The positive relationship between problem-solving ability (as measured in practice on authentic tasks using the problem-solving assessment instrument) and educational achievement (as indicated by the grade achieved by students on the module in which the problem-solving assessment was conducted) supports the first hypothesis. Path models, such as the one shown in figure 9, generally represent causal relationships. However, given the simple structure of this model, in which the two latent constructs PSAbil and EdAch each represented only one manifest variable, it would be unwise to assert causality. A much more comprehensive model that takes into account prior educational achievement is required.

The lack of any relationship between attitude and educational achievement is quite surprising, especially given that higher grades are awarded only to those who are motivated to seek assessment against additional criteria and who demonstrate the corresponding competencies. In other words, this form of assessment incorporates a motivational component, and it is expected that the motivation to seek higher levels of assessment will be related strongly to favourable attitudes about an individual's problem-solving confidence and control.

The negative relationship between attitude and problem-solving ability is equally surprising. This result is counter to models of domain learning such as the one proposed by Murphy and Alexander (2002).

In order to explain the surprising relationships between attitude and both educational achievement and problem-solving ability, it will be necessary to revisit the problem-solving inventory. The current context did provide a situation in which the predictive validity of the problem-solving inventory could be tested. The failure to observe expected relationships must cast some doubt on the predictive validity of the problem-solving inventory. However, it was not the purpose of this study to test that instrument. The primary purpose of the project was to test a novel approach to the assessment of problem-solving ability, and the strong positive relationship between this ability as measured by the problem-solving assessment instrument and educational achievement provides an indication of the concurrent validity of the instrument.

Evaluation by students

An evaluation by students of the key competencies assessment process employed in the Authentic Performance-based Assessment of Problem-Solving study showed that students were quite clear about the benefits to them of participating in key competencies assessments. They understood that through these assessments their skill levels would receive formal recognition and that this would assist them in seeking employment. In addition, they understood that through the processes employed in which they self-assessed their problem-solving ability they would enhance both their key competencies and their abilities to talk about them.

There was some support for the more detailed approach taken in administering the problem-solving assessment instrument compared with previous tools. However, there is a need to review the terminology used in the instrument to make it more accessible to students.

The concerns that students expressed about performance levels suggest the need for further work to be done on this. Performance levels have been a difficulty from their initial articulation in the Mayer report. In the consultations undertaken by the Mayer Committee, this issue was identified as a problem. Since that time several authors have commented on confusion surrounding these levels. In the current study, the issue that taxed students was the clarity with which the different levels were distinguished. The support expressed by students for the more detailed indicators of the problem-solving assessment instrument compared with previous assessment tools suggests that this might be an effective approach to the resolution of this matter.

6 Conclusions and future directions

In this chapter, discussion is restricted to the major issues that have arisen from the study. The Authentic Performance-based Assessment of Problem-Solving Project had two principal aims, expressed as guiding questions:

- Is it feasible to extend existing problem-solving tasks through the development of scoring rubrics that address problem-solving processes in order to develop reliable assessments of problem-solving ability?
- Is there an identifiable positive relationship between indicators of problem-solving ability and subsequent learning and use of knowledge about problem-solving tasks?

The purpose of the first of these questions was to ascertain whether a valid and reliable general-purpose problem-solving instrument could be developed and deployed to assess problem-solving ability in the tasks that students routinely undertake in their courses.

The second question subsumes several propositions. First, if problem-solving is an important attribute, its processes should lead to successful learning. Thus evidence of an association between problem-solving performance and course success should be forthcoming. Second, the explicit attention learners paid to problem-solving processes was expected to result in greater knowledge about problem-solving and to greater control over problem-solving in practice. That is, the more often students engaged in the assessment of their problem-solving, the greater their problem-solving performance should have been. Third, the second question implies a test of the transfer of problem-solving skills across tasks, initially within the domain of the student's course.

The study had several limitations. These include:

- The Authentic Performance-based Assessment of Problem-Solving Project was implemented in a specific field of study in one TAFE institute.
- The students who participated were volunteers and may not be typical of all students in their course.
- The study design required large numbers of participants in order to establish statistical significance for some of the hypothesised outcomes and, although 43 usable assessments were available in the test sample, these were insufficient to permit the multilevel analysis that had been envisaged.

Despite these limitations, some conclusions can be reached on the basis of the data that have been analysed. In response to these limitations and to some of the findings that have emerged from the study, suggestions are made both for the further use of the problem-solving assessment instrument and for additional research that might be undertaken to overcome the identified limitations.

The feasibility of employing a general-purpose problem-solving assessment

The problem-solving assessment instrument was developed following an extensive review of major theories of problem-solving. A convergence of theoretical positions on what constituted problem-

solving was found and provided a basis for the development of an instrument that can claim a high level of construct and content validity.

The instrument was trialled within the Electronics and Information Technology Program at Torrens Valley Institute of TAFE. It was used by seven staff members in the assessment of 33 students (the test sample) who collectively had 43 tasks assessed for their problem-solving performance. In all, 25 tasks had been identified as potential targets for problem-solving assessments. Of these, 18 were the subject of assessments.

The problem-solving assessment instrument was also tested in a validation study within the Certificate IV in Assessment and Workplace Training with a sample of 48 respondents.

Analyses of the results obtained from the administration of the problem-solving assessment instrument revealed that it has generally very good psychometric properties and therefore that it meets criteria for reliable assessment. However, some differences between the original test sample and the validation sample point to the need to revise performance criteria for some indicators. The generally good psychometric properties of the current version of the problem-solving assessment instrument, coupled with the information provided by the various analytical techniques used in its examination, should enable a very robust and reliable instrument to be developed.

Future development of the problem-solving assessment instrument

The problem-solving assessment instrument represents an advance over previous research efforts to assess problem-solving. Instrumental assessment has been shown to be reliable; for example, in the graduate skills assessment (Australian Council for Educational Research 2001b), but these approaches are criticised for lacking authenticity. Authentic assessment using 'real' tasks (for example, Herl et al. 1999) has two particular problems. First, a separate scoring rubric has been required for each task and this involves a very heavy testing overhead. Second, these specific approaches lead to considerable variability in performance that is related to each task, which in turn requires that performance on many tasks be undertaken by each individual in order to achieve a reasonable level of reliability. The generality of the problem-solving assessment instrument certainly overcomes the first of these problems and makes inroads into the second, although more research is required to demonstrate the extent to which assessment is task independent.

Assessments based on the instrument are of known precision. That precision will form a basis for making judgments about the number of levels of problem-solving performance that can be discriminated within a sample of assessments. It appears that at least three bands of performance can be recognised among successful candidates. The inclusion of greater numbers of less successful performances might lead to the identification of additional bands of performance.

The instrument does need to be refined. What is not clear is whether the problem-solving assessment instrument is appropriately targeted because the sample of respondents were volunteers and there may have been a sampling bias in favour of higher ability students. The targeting of the instrument was reasonable, but could be improved for the current sample of respondents if more rigorous indicators were added. The responses of the validation study group also suggest the need for more rigorous indicators. Alternatively, some of the existing indicators could be made more challenging if their performance levels and criteria were revised to make them more difficult to achieve.

Also not yet clear is the extent to which the instrument would work reliably in other domains. It has been shown to work well in one technical domain and it has been shown to work in one other field of study. However, some systematic differences between the two groups suggest a need to revise performance criteria for some indicators. What remains is for the instrument to be tested in other technical areas and in non-technical fields.

The development of problem-solving ability

An issue that has been of concern to policy-makers and practitioners is whether key competencies and other generic skills are both teachable and assessable. This study has shown that problem-solving is certainly assessable using a general-purpose assessment tool. The two-stage method employed in this study—self-assessment followed by lecturer validation, with both the learner and the assessor using the same instrument—has provided an indication that problem-solving skill is enhanced through this process; that is, the assessment process was also a learning process for participants.

While it is certainly possible to use the problem-solving assessment instrument as a summative assessment tool, the assessment process used in the Authentic Performance-based Assessment of Problem-Solving Project demonstrates that assessment exercises can also be learning activities for students and can therefore increase the efficiency of the learning and assessment process. In this project, students were taught vocational competencies and were tested on them and, in association with that process, were also assessed on a key competency (problem-solving) in a way that contributed to the development of that competency.

The relationships between attitude, problem-solving ability and educational achievement

Performance within a domain is related strongly to prior achievement, the extent and organisation of individuals' knowledge, and individual motivation. There are also contextual factors related to the learning environment and these include the quality of instruction and peer interaction and support. Among motivational factors, interest in the topic and perceived self-efficacy have been identified as contributors to performance. In this study, the role of problem-solving ability as measured by the problem-solving assessment instrument was added and a simple model relating attitude measures, problem-solving ability and educational achievement was tested.

A very strong relationship was found between problem-solving ability and educational achievement. The model proposed in this study is too simple to be able to claim a causal link from problem-solving ability to educational achievement. The association between these constructs may be related to individual ability. However, the indication that the assessment process led to the enhancement of this ability, and the association between problem-solving ability and educational achievement, suggest that the specific and contextual assessment of this capacity is a worthy endeavour.

Future directions

Both the favourable outcomes of the study and some of the limitations that have attended the Authentic Performance-based Assessment of Problem-Solving Project suggest opportunities for the extension of this work. The instrument needs to be trialled with a broader cross-section of candidates, in terms of their abilities, backgrounds and fields of study. The main study and the validation study have provided information that can be used in the revision of some indicators. On the basis of more extensive trials, further revisions of the problem-solving assessment instrument might be undertaken.

The methodology of the development of the instrument suggests that comparable instruments for the assessment of other generic skills could be developed using this approach. The processes employed in its implementation in this project suggest that similar approaches could be used in its wider implementation as a means for promoting the development of problem-solving ability in other contexts.

Further implementation of the problem-solving assessment instrument

The problem-solving assessment instrument has been shown to be a valid and reliable tool for the assessment of problem-solving ability within a restricted range of programs. In order to ascertain whether the tool is as generally applicable as it was intended to be, it must be trialled in a wide range of courses and settings.

First, it would seem sensible to test the problem-solving assessment instrument in courses offered by a range of providers in the VET sector. A wide range of courses should also be targeted. Priority in the selection of courses might be based on the number of participants engaging in a field of study and on the nature of the courses within it. The instrument has been tested in a technical program and validated in one non-technical course. It might now be trialled in additional technical courses as well as in business, personal service and human services programs.

Second, the problem-solving assessment instrument has been used in a course that services full- and part-time students who attend their courses on campus. It is desirable to trial the instrument in off-campus, distance-delivered programs and in workplace-based learning situations.

Finally, a comparable form of assessment is required to enable people who have completed courses and who want access to the certification of their abilities through recognition of current competency arrangements.

It is suggested that the problem-solving assessment instrument be trialled in a range of fields of study and settings in order to establish its generality as an assessment tool.

Refinement of the problem-solving assessment instrument

One of the limitations of the Authentic Performance-based Assessment of Problem-Solving Project was the modest number of participants who were assessed using the problem-solving assessment instrument. The number was sufficient to validate the structure of the instrument and to provide useful indications of the qualities of its indicators. However, in order to undertake meaningful refinements to the instrument, it does need to be tested with a greater number of participants.

The suggestion made above to trial the instrument with a broader range of participants in diverse settings will go some way to addressing the number of participants. However, as new courses and settings were added, new variables would also be added to the analyses, and each new variable would require an increase in the number of participants in order to establish the statistical significance required to either support or negate propositions about the instrument's measurement properties.

Although favourable indications have been found, the robust demonstration of claims that the application of the instrument leads to enhanced problem-solving abilities among participants requires multilevel analytical techniques. For these to work, the instrument needs to be tested in a variety of courses and settings, with a large number of individuals, and on at least three occasions for each individual.

A further dimension of the assessment of problem-solving was canvassed in the review of literature. Shavelson et al. (1993) noted that task variability contributed more to variation in assessed performance than did inter-rater variation. In this study, an attempt was made to moderate the influence of task variability, but several facets might have contributed to variation in assessed performance. Facets that warrant attention include task, occasion and rater. In order to identify the influences of these variables, a substantial number of participants would be required.

It is suggested that a large-scale trial of the problem-solving assessment instrument be undertaken across a range of providers and in a range of courses engaging several hundred participants, each of whom is assessed on at least three occasions over the course of a calendar year.

As part of such a trial, it will be desirable to establish the concurrent and predictive validity of the problem-solving assessment instrument. This can be done by gathering data on prior educational achievement of participants and on their performance on related measures of problem-solving ability. The adaptation of instruments such as the graduate skills assessment might be considered for this purpose.

Procedures for key competencies assessments

Before the trial of the problem-solving assessment instrument, staff in the Electronics and Information Technology Program at Torrens Valley Institute had encouraged students to seek key competencies assessments by selecting existing assignment tasks from within the modules they were taking.

Within the Authentic Performance-based Assessment of Problem-Solving Project, a subset of existing assignment tasks was nominated by staff, and students were encouraged to seek a problem-solving assessment of them using the problem-solving assessment instrument. Staff nominated these tasks because they were thought to represent productive opportunities for students to demonstrate their problem-solving skills. Some tasks were selected from introductory modules and others from more advanced ones. Students were also able to submit for assessment work that they had undertaken on tasks of their own choice.

In informal discussions with staff in other provider institutions, who were not involved in the project, it emerged that the requirement to assess each key competency in each unit of competency was perceived to be an onerous assessment load for both staff and students. This load was a factor in staff's not being willing to employ an assessment tool that they felt would add to the demands on their time.

The approach of requiring key competencies assessments on a selection of tasks may overcome the perceived assessment load barrier and lead to robust and reliable assessments and subsequent reporting of performance on these valued skills.

It is suggested that the assessment processes documented in the Authentic Performance-based Assessment of Problem-Solving Project be trialled by other providers with the intention that developers of training packages consider the adoption of this approach to the assessment of key competencies in future revisions of those packages.

The assessment procedure that was adopted in the project, which included a self-assessment component, was:

- Staff nominated a set of existing assessment tasks within units of competency that they believed would give students the opportunity to demonstrate their skills.
- For each key competency, staff nominated a set of relevant tasks.
- Participating students selected and undertook any three of the tasks that were listed for each key competency and submitted their work on those tasks as evidence of the achievement of the corresponding key competency.
- Students used an evidence guide to assist them in presenting a case for a particular level of performance on each key competency (their self-assessment of their problem-solving ability).
- Staff used an assessment tool, incorporating evidence criteria, to support the judgements they made of the evidence students presented to them.

The development of other assessment tools

The process that was followed in the development of the problem-solving assessment instrument is:

- Theoretical conceptions of problem-solving were explored in order to arrive at a coherent description of problem-solving. This provided a basis for the claim of construct validity for the concept as it was implemented.
- Major component processes in problem-solving were identified from a variety of theoretical positions. These major processes served to delineate the scope of problem-solving and to establish a basis for the content validity of this implementation.
- For each of the five major processes identified, a set of indicators was proposed. These indicators operationalised the major processes and linked the theoretical foundation of the problem-solving assessment instrument to its practical implementation.
- Finally, for each indicator, a set of performance levels was described. The basis for setting performance levels for indicators was the SOLO taxonomy. This provided a sound reference scale for performance levels on diverse indicators. These levels provided a basis for scoring the evidence that learners presented to support their claims of the use of problem-solving processes.

This method resulted in the development of an instrument that is both valid and reliable. However, only one key competency (problem-solving) has been assessed with an instrument that reflects this genesis.

It is suggested that this development method be followed in the production of instruments to assess and measure other key competencies and that similar validation processes be pursued in the evaluation of those instruments.

Levels of performance

The issue of performance levels has been a continuing challenge in the VET sector. The Mayer Committee found this in their consultations with industry (Mayer Committee 1992, appendix 3) and it has been reported in a number of studies since that time.

The Rasch analysis of the problem-solving assessment results does suggest that the instrument may provide a basis for identifying more than three levels of performance and, on the basis of the current data set, it seems that five performance bands could be discriminated. An advantage of the structure of the problem-solving assessment instrument is that the individual indicators do not all need to reflect the proposed overall performance levels. Some items in the problem-solving assessment instrument are dichotomous while others may have four or five performance levels. Clearly, the availability of this powerful measurement model does provide a basis for establishing objectively both the number of performance levels and cut scores to separate levels. Achieving this goal in practice will require more extensive studies, such as those recommended above.

It is suggested that, in the wider application of the problem-solving assessment instrument, analyses be undertaken to establish both the number of performance levels that can be discriminated and the criteria by which performance levels can be established.

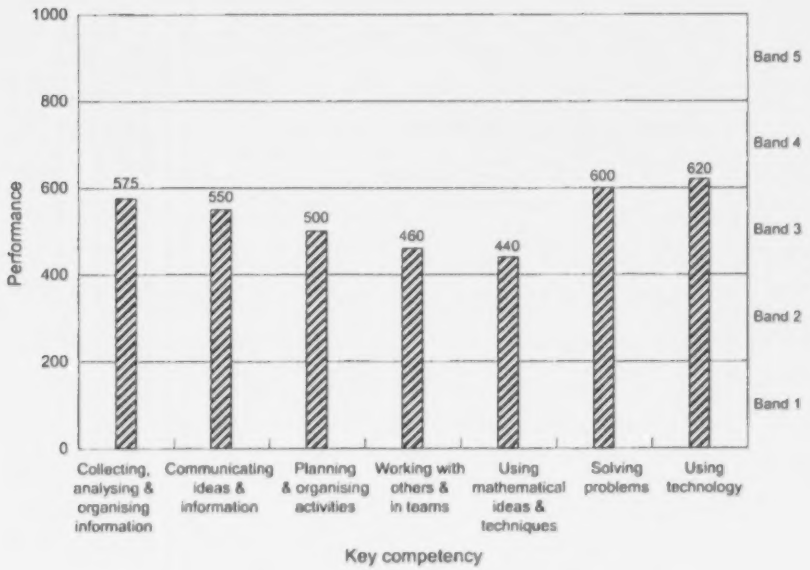
Forms of reporting for assessments of key competencies

If comparable assessments of other key competencies can be developed and their results analysed using the tools that have been employed in this study, it will be possible to generate reports that are based on robust assessment processes that produce ability estimates of known precision on comparable scales.

The generation of ability estimates of known precision seems to be a highly desirable goal for training providers. Apart from adding value to the vocational dimension of the education and training that such estimates deliver, reporting based on demonstrably robust and reliable processes should enhance the confidence that both employers and graduates of VET providers can have in the assessment of these important skill sets.

It is becoming apparent that, while key competencies are indeed important and valued across industry, they operate in clusters, and that strengths in particular combinations of key competencies might be highly valued in particular industries or for particular jobs. For example, a person who will work in a 'back-room' capacity will need, in addition to technical skills, well-developed problem-solving and teamwork skills. A colleague in the same enterprise who works in customer service may require a different profile of skills—one that emphasises problem-solving and communication abilities. A form of reporting that reveals key competencies profiles would be advantageous. An example, based on simulated performance data, is shown in figure 10. Such a report, with information on the meanings of performance bands, may be informative for learners and potential employers. To be a valuable record of achievement, it would need to be based on robust instruments and assessment procedures.

Figure 10: Simulated key competencies profile for inclusion in a report



It is suggested that the methods employed in the Authentic Performance-based Assessment of Problem-Solving Project be applied to the development of assessment instruments for other key competencies and that reporting methods based on the application of robust measurement methods be developed.

References

Adams, RJ & Khoo, ST 1999, Quest: The interactive test analysis system (version PISA) [statistical analysis software], Australian Council for Educational Research, Melbourne.

Airasian, PW 1994, *Classroom assessment* (2nd edn), McGraw Hill, New York.

Alexander, PA & Judy, JE 1988, 'The interaction of domain-specific and strategic knowledge in academic performance', *Review of Educational Research*, vol.58, no.4, pp.375–404.

Allen Consulting Group 1999, *Training to compete: The training needs of Australian industry: A report to the Australian Industry Group*, Australian Industry Group, North Sydney.

American Educational Research Association, American Psychological Association & National Council on Measurement in Education 1985, *Standards for educational and psychological testing*, American Psychological Association, Washington DC.

Anderson, LW 1997, 'Attitudes, measurement of', in *Educational research, methodology, and measurement: An international handbook*, ed. JP Keeves, Pergamon, Oxford, pp.885–895.

Anderson, LW & Krathwohl, DR (eds) 2001, *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (abridged edn), Addison-Wesley Longman, New York.

Australian Chamber of Commerce and Industry & Business Council of Australia 2002, *Employability skills for the future*, Department of Education, Science and Training, Canberra.

Australian Council for Educational Research 2000, *Graduate skills assessment: Test development and progress report*, ACER, Camberwell, Victoria.

—— 2001a, *Graduate skills assessment: Summary report*, DETYA, Higher Education Division, Canberra.

—— 2001b, *Graduate skills assessment: Summary report: GSA exit 2000*, ACER, Camberwell, Victoria.

Australian Vice Chancellors Committee 1997, *Joint NHMRC/AVCC statement and guidelines on research practice*, Australian Vice Chancellors Committee, Canberra.

Bailey, T 1997, 'Changes in the nature of work: Implications for skills and assessment', in *Workforce readiness: Competencies and assessment*, ed. HF O'Neil, Lawrence Erlbaum and Associates, Mahwah, New Jersey, pp.27–45.

Bandura, A 1989, 'Regulation of cognitive processes through perceived self-efficacy', *Developmental Psychology*, vol.25, no.5, pp.729–735.

Bandura, A 1997, *Self-efficacy: The exercise of control*, WH Freeman, New York.

Biggs, JB & Collis, KF 1982, *Evaluating the quality of learning: The SOLO taxonomy (Structure of the Observed Learning Outcome)*, Academic Press, New York.

Bond, TG & Fox, CM 2001, *Applying the Rasch model: Fundamental measurement in the human sciences*, Lawrence Erlbaum and Associates, Mahwah, New Jersey.

Bransford, JD & Stein, BS 1984, *The ideal problem solver: A guide for improving thinking, learning, and creativity*, WH Freeman, New York.

Bryk, A, Raudenbush, S & Congdon, R 1996, HLM for Windows: Hierarchical linear and nonlinear modeling with HLM/2L and HLM/3L (version 4) [Multilevel analysis software], Scientific Software International, Chicago.

Camp, CJ 1992, 'The problem solving inventory (test 303)', in *Eleventh mental measurements yearbook*, eds JJ Kramer & JC Conoly, Gryphon Press, Highland Park, New Jersey, pp.699–701.

Conference Board of Canada 2000, *Employability skills toolkit for the self-managing learner* [CD-ROM kit], Ryerson McGraw-Hill, Canada, <http://www.conferenceboard.ca/education/learning-tools/toolkit.htm> [viewed 4 February 2001].

Curtis, DD 1996, 'Using technology', in *Teaching and learning the key competencies in the vocational education and training sector: Research support*, eds DD Curtis, R Hattam, JP Keeves, MJ Lawson, R Reynolds, A Russell, H Silins & J Smyth, Flinders Institute for the Study of Teaching, Adelaide.

—— 2001, 'The problems of assessing problem-solving', unpublished report, Flinders University, Adelaide.

Curtis, DD, Hattam, R, Keeves, JP, Lawson, MJ, Reynolds, R, Russell, A, Silins, H & Smyth, J (eds) 1996, *Teaching and learning the key competencies in the vocational education and training sector: Research support*, Flinders Institute for the Study of Teaching, Adelaide.

Curtis, DD & McKenzie, P 2002, *Employability skills for Australian industry: Literature review and framework development*, Department of Education, Science and Training, Canberra.

BEST COPY AVAILABLE

Dawe, S 2001, 'Do training packages focus sufficiently on generic skills?', paper presented at *Knowledge Demands for the New Economy,* Ninth Annual International Conference on Post-compulsory Education and Training, 3–5 December, Surfers Paradise, Queensland.

Down, CM 2000, 'Key competencies in training packages', paper presented at *Learning Together, Working Together: Building Communities in the 21st Century,* 8th Annual International Conference on Post-Compulsory Education and Training, 4–6 December, Gold Coast, Queensland.

Dweck, CS 1986, 'Motivational processes affecting learning', *American Psychologist,* vol.41, no.10, pp.1040–1048.

Employment and Skills Formation Council 1992, *The Australian vocational certificate training system* (Carmichael report), National Board of Employment, Education and Training, Canberra.

Feast, V 2000, 'Student perceptions of the importance and value of a graduate quality framework in a tertiary environment', unpublished Doctor of Education dissertation, Flinders University, Adelaide.

Field, L 2001, *Employability skills required by Australian workplaces,* Field Learning, Sydney.

Finn Review Committee 1991, *Young people's participation in post-compulsory education and training: Report of the Australian Education Council Review Committee,* Australian Government Publishing Service, Canberra.

Gillis, S & Bateman, A 1999, *Assessing in VET: Issues of reliability and validity,* National Centre for Vocational Education Research, Adelaide.

Greeno, JG, Collins, AM & Resnick, LB 1996, 'Cognition and learning', in *Handbook of educational psychology,* eds DC Berliner & RC Calfee, Macmillan, New York, pp.15–46.

Griffin, P 2000, 'Competency based assessment of higher order competencies', keynote address presented at the NSW ACEA State Conference, Mudgee, 28 April 2000, <http://www.edfac.unimelb.edu.au/LED/ARC/PDFs/aceansw.pdf> [viewed 28 June 2001].

—— 2001, 'Performance assessment of higher order thinking', paper presented at the annual conference of the American Educational Research Association, Seattle, 10 April 2001, <http://www.edfac.unimelb.edu.au/LED/ARC/PDFs/griffin2001.pdf> [viewed 28 June 2001].

Griffin, P, Gillis, S, Keating, J & Fennessy, D 2001, *Assessment and reporting of VET courses within senior secondary certificates,* commissioned report, NSW Board of Vocational Education and Training, Sydney.

Hager, P & Beckett, D 1999, 'Making judgments as the basis for workplace learning: Preliminary research findings', proceedings of *Quality and Diversity in VET Research,* the Australian Vocational Education and Research Association (AVETRA) Conference, 11–12 February 1999, RMIT, Melbourne, <http://www.avetra.org.au/abstracts-papers/19_hager.pdf> [viewed 23 December 2000].

Hamlbur, S & Glickman, H 2001, *Summary report: GSA exit 2000,* ACER, Melbourne.

Harwell, MR & Gatti, GG 2001, 'Rescaling ordinal data to interval data in educational research', *Review of Educational Research,* vol.71, no.1, pp.105–131.

Hase, S 2000, 'Measuring organisational capability: Beyond competence', proceedings of *Future Research, Reseach Futures,* the Australian Vocational Education and Research Association (AVETRA) Conference, 23–24 March, Canberra, <http://www.avetra.org.au/papers%202000/hase.PDF> [viewed 7 June 2001].

Heppner, PP & Petersen, CH 1982, 'The development and implications of a personal problem-solving inventory', *Journal of Counseling Psychology,* vol.29, no.1, pp.66–75.

Herl, HE, O'Neil, HF, Chung, GKWK, Bianchi, C, Wang, S-L, Mayer, R, Lee, CY, Choi, A, Suen, T & Tu, A 1999, *Final report for validation of problem-solving measures,* CSE technical report 501, Center for the Study of Evaluation and National Centre for Research in Evaluation, Standards, and Student Testing, Los Angeles.

Jasinski, M 1996, *Teaching and learning the key competencies in vocational education and training,* Western Adelaide Institute of TAFE, Port Adelaide, South Australia.

Kearns, P 2001, *Generic skills for the new economy: A review of research relating to generic skills,* National Centre for Vocational Education Research, Adelaide.

Keeves, JP & Kotte, D 1996, 'The measurement and reporting of key competencies', in *Teaching and learning the key competencies in the vocational education and training sector: Research support,* eds DD Curtis, R Hattam, JP Keeves, MJ Lawson, R Reynolds, A Russell, H Silins & J Smyth, Flinders Institute for the Study of Teaching, Adelaide, pp.139–167.

Kotovsky, K, Hayes, JR & Simon, HA 1985, 'Why are some problems hard? Evidence from Tower of Hanoi', *Cognitive Psychology,* January, pp.248–294.

Lave, J 1988, *Cognition in practice: Mind, mathematics and culture in everyday life,* Cambridge University Press, Cambridge, New York.

Lawson, MJ & Hopkins, S 1996, 'Solving problems in commercial cookery', in *Teaching and learning the key competencies in the vocational education and training sector: Research support,* eds DD Curtis, R Hattam, JP Keeves, MJ Lawson, R Reynolds, A Russell, H Silins & J Smyth, Flinders Institute for the Study of Teaching, Adelaide, pp.17–43.

Lokan, J (ed.) 1997, *Describing learning: Implementation of curriculum profiles in Australian schools 1986–1996,* ACER Press, Camberwell, Victoria.

Masters, G & Forster, M 2000, 'The assessments we need', unpublished paper, Australian Council for Educational Research, Camberwell, <http://www.acer.edu.au> [viewed 8 September 2000].

BEST COPY AVAILABLE

Mayer Committee 1992, *Key competencies: Report of the Committee to advise the Australian Education Council and Ministers of Vocational Education, Employment and Training on employment-related key competencies for postcompulsory education and training*, Australian Education Council and Ministers of Vocational Education, Employment and Training, Canberra.

Mayer, JD 2001, 'Emotion, intelligence, and emotional intelligence', in *Handbook of affect and social cognition*, ed. JP Forgas, Lawrence Erlbaum and Associates, Mahwah, New Jersey, pp.410–431.

Mayer, RE 1992, *Thinking, problem solving, cognition* (2nd edn), WH Freeman, New York.

Mayer, RE & Wittrock, MC 1996, 'Problem-solving transfer', in *Handbook of educational psychology*, eds DC Berliner & RC Calfee, Macmillan, New York, pp.47–62.

McCurry, D & Bryce, J 1997, *The school-based key competencies levels assessment project: Final report*, DEETYA, Canberra.

—— 2000, *Victorian Board of Studies: Key competencies levels assessment trial: Working paper 2*, Victorian Curriculum and Assessment Authority, Melbourne.

McGaw, B 1997, *Shaping their future: Recommendations for the reform of the higher school certificate*, Department of Training and Education Co-ordination, Sydney.

Michell, J 1997, 'Quantitative science and the definition of measurement in psychology', *British Journal of Psychology*, vol.88, pp.355–383.

Murphy, PK & Alexander, PA 2002, 'What counts? The predictive powers of subject-matter knowledge, strategic processing, and interest in domain-specific performance', *Journal of Experimental Education*, vol.70, no.3, pp.197–214.

National Industry Education Forum 2000, *The key competencies portfolio approach: A kit*, Department of Education, Training and Youth Affairs, Canberra.

Newell, A & Simon, HA 1972, *Human problem solving*, Prentice Hall, Englewood Cliffs.

OECD (Organisation for Economic Co-operation and Development) & Statistics Canada 1995, *Literacy, economy and society: Results of the First International Adult Literacy Survey*, OECD and Statistics Canada, Paris.

O'Keefe, S 2000, 'Management education: A case study: Student perceptions of the role of the employer in formal off-the-job education', proceedings of *Future Research, Reseach Futures*, the Australian Vocational Education and Research Association (AVETRA) Conference, 23–24 March, Canberra, <http://www.avetra.org.au/papers%202000/okeefe.PDF> [viewed 7 June 2001].

Osterlind, SJ 1998, *Constructing test items: Multiple choice, constructed response, performance, and other formats* (2nd edn), Kluwer Academic Publishers, Boston.

Pellegrino, J, Chudowsky, N & Glaser, R (eds) 2001, *Knowing what students know: The science and design of educational assessment: A report of the National Research Council*, National Academy Press, Washington DC.

Polya, G 1957, *How to solve it: A new aspect of mathematical method* (2nd edn), Penguin, Harmondsworth.

Quality of Education Review Committee 1985, *Quality of education in Australia: Report of the Review Committee*, Australian Government Publishing Service, Canberra.

Queensland Department of Education 1997, *Assessing and reporting the key competencies of students of post-compulsory age through 'work experience'*, Department of Education, Employment, Training and Youth Affairs, Canberra.

Reynolds, C 1996, *Business, industry, key competencies and portfolios*, Department of Education, Employment, Training and Youth Affairs, Canberra.

Reynolds, R & van Eyk, B 1996, 'Training and key competencies: Other points', in *Teaching and learning the key competencies in the vocational education and training sector: Research support*, eds DD Curtis, R Hattam, JP Keeves, MJ Lawson, R Reynolds, A Russell, H Silins & J Smyth, Flinders Institute for the Study of Teaching, Adelaide.

Robertson, I, Harford, M, Strickland, A, Simons, M & Harris, R 2000, 'Learning and assessment issues in apprenticeships and traineeships', proceedings of *Future Research, Reseach Futures*, the Australian Vocational Education and Research Association (AVETRA) Conference, 23–24 March, Canberra, <http://www.avetra.org.au/papers%202000/robetal.PDF> [viewed 7 June 2001].

Russell, A 1996, 'Working with others and in teams', in *Teaching and learning the key competencies in the vocational education and training sector: Research support*, eds DD Curtis, R Hattam, JP Keeves, MJ Lawson, R Reynolds, A Russell, H Silins & J Smyth, Flinders Institute for the Study of Teaching, Adelaide, pp.45–76.

Salovey, P, Mayer, JD, Goldman, SL, Turvey, C & Palfai, TP 1995, 'Emotional attention, clarity, and repair: Exploring emotional intelligence using the trait meta–mood scale' in *Emotion, disclosure, & health*, ed. JW Pennebaker, American Psychological Association, Washington DC, pp.125–154.

Scribner, S 1986, 'Thinking in action: Some characteristics of practical thought', in *Practical intelligence: Nature and origins of competence in the everyday world*, eds RJ Sternberg & RK Wagner, Cambridge University Press, Cambridge, pp.13–30.

Sellin, N 1989, *PLS Path* (version 3.01) [statistical analysis software], N Sellin, Hamburg.

BEST COPY AVAILABLE

Shavelson, RJ, Gao, X & Baxter, GP 1993, *Sampling variability of performance assessments*, CSE technical report 361, University of California, CRESST, Los Angeles.
Sheridan, B, Andrich, D & Luo, G 1997, RUMM (version 2.7Q) [statistical analysis software], RUMM Laboratory, Perth.
Smith, E 2000, 'Teenagers' full-time work and learning: A case study in what research findings say about policy, practice, theory and further research possibilities', proceedings of *Future Research, Research Futures*, the 2000 Australian Vocational Education and Research Association (AVETRA) Conference, 23–24 March, Canberra, <http://www.avetra.org.au/papers%202000/smithe.PDF> [viewed 7 June 2001].
SPSS Inc. 1995, SPSS for Windows (version 6.1.3) [statistical analysis program], SPSS Inc. Chicago.
—— 2000, SPSS for Windows (version 10.0.7) [statistical analysis program], SPSS Inc. Chicago.
Sternberg, RJ 1985, *Beyond IQ: A triarchic theory of human intelligence*, Cambridge University Press, London.
Troper, J & Smith, C 1997, 'Workplace readiness portfolios', in *Workforce readiness: Competencies and assessment*, ed. HF O'Neil, Lawrence Erlbaum and Associates, Mahwah, New Jersey, pp.357–382.
Wiggins, G 1989, 'A true test: Toward more authentic and equitable assessment', *Phi Delta Kappan*, vol.70, no.9, pp.703–713.
Wright, BD & Masters, G 1981, *The measurement of knowledge and attitude, research memorandum 30*, University of Chicago, Department of Education, Statistical Laboratory, Chicago.
Zeller, RA 1997, 'Validity', in *Educational research, methodology, and measurement: An international handbook*, ed. JP Keeves, Pergamon, Oxford, pp.822–829.

BEST COPY AVAILABLE

Appendices

Appendix 1: Consent and personal details forms

Participation consent form

This project, the Authentic Performance-Based Assessment of Problem-Solving, is an attempt to find a new way to assess and report on individual's problem-solving skills in real-world situations. Problem-solving skills are to be assessed as part of routine course assessments.

I hereby give my consent to David D Curtis, a researcher in the Centre For Lifelong Learning and Development and the School of Education at Flinders University, and whose signature appears below, to access my problem-solving assessment summaries.

I understand that my participation in the research will include:

- Completion of a form outlining some individual details (name, age, gender, prior education, and work background);
- Completion of a Problem-Solving Inventory questionnaire; and
- The assessment of at least one, and up to three, of the normal course assessment tasks for their problem-solving aspects.

I also understand that, for the purposes of this research, the grades obtained in my studies to date can be accessed for use in this research project.

I give permission for the use of these data, and of other information which I have agreed may be obtained or requested, in the writing up of the study, subject to the following condition:

- That my identity and the results that I achieve in the assessment remain confidential.

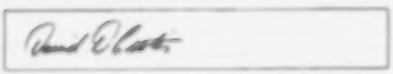
I understand that I will be given a report on the Problem-Solving Inventory results and on the Problem-Solving Assessments for each of the tasks submitted. I also understand that I will be granted an additional day on the expiry date for each course module for which a Key Competency problem-solving assessment is submitted.

My participation in this study is voluntary, and I understand that I may withdraw from the study at any time.

Signatures

Researcher

Name David D Curtis

Signature  Date 11 March 2002

Participant

Name

Signature _____ Date

Personal details form

The Authentic Performance-Based Assessment of Problem-Solving

Personal Details Form

Name:

Age (years)

Gender

- ❑ Male
- ❑ Female

Current Course

Educational Experience

School Education

(Please indicate the highest level of school education completed)

Less than Year 10	❑
Year 10	❑
Year 11	❑
Year 12	❑

Post-School Education

(Please indicate the highest level of post-school education completed)

No post-school education other than current course	❑
Partial completion of another TAFE course	❑
Completion of another TAFE course	❑
Completion of a university course	❑

Work Experience

Please indicate the number of full time years of work experience in any field, including unpaid work

No previous work experience	❑
Up to 1 year	❑
1 to 5 years	❑
More than 5 years	❑

Appendix 2: The problem-solving assessment procedure

Electronics & Information Technology
Problem Solving Project

Assessment Procedure

1. COMPLETE initial project documentation

- Consent Form
- Personal Details
- Problem Solving Inventory

To be done ONCE only – should take less than 20 minutes total – forward all forms to Rob

2. CHOOSE ANY appropriate module assessment

You can CHOOSE ANY **module assessment** to demonstrate your Problem Solving skills.
A list of '**recommended**' module assessment activities has been prepared to help you.

Recommended Assessment Activities

MODULE	ASSESSMENT ACTIVITY
AC Principles	Practical Test
Applied Electricity 1	Practical Activity 2.1 – Electrical Circuits
Digital Electronics 1	Practical Activity 3.1 – Digital Displays
Digital Subsystems 1	Practical Activity 2.2 – Switch Debouncing
Electrical Fundamentals	Prac Test 1
Electrical Fundamentals	Destructive Activity
Embedded Systems	Practical Activity 2.1 – Interfacing a Microcontroller Development Board to a PC
Hardcopy	Practical Activity 3.1.1 – Practical Application
Intro to Electricity & Electronics	Practical Activity 2.2 – Basic Circuit Measurements
Intro to Electricity & Electronics	Practical Activity 4.1 – Destructive action
Intro to Programming	Final Programming Project – Game Program (3-in-a-row)
Microprocessor Fundamentals	Practical Activity 3.1 – Using a RAM Chip
PC Systems Support	Practical Activity 6.1 – Application/Hardware Installation
Power Supply Principles	Topic 11 – Fault-finding Techniques
Single User Operating Systems	Topic 4.1 – Configuring Operating System
Single User Operating Systems	Crash Activity
Soldering – Basic	Project kit

3. SELF ASSESSMENT

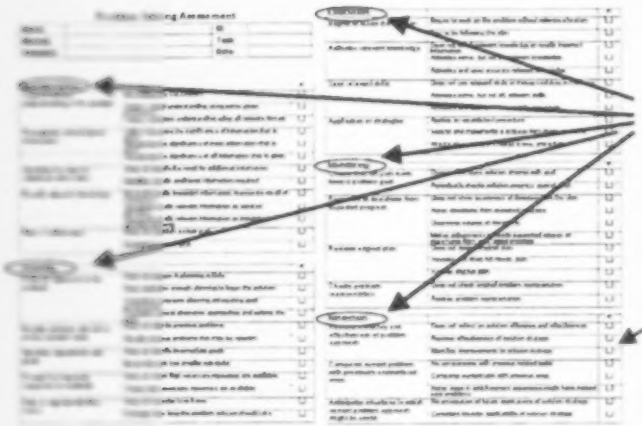
Use the ***Problem Solving Assessment*** sheet to
(1) Self assess your performance AND
(2) Identify evidence for EACH item ticked

- You need to be able to point out or briefly explain **how** you addressed each item ticked
- Evidence may be written, demonstrated, discussed or pointed out in assessment reports etc

Important Note: *your evidence must highlight the* ***Problem Solving processes*** *you used (rather than your technical knowledge)*

See over page →

BEST COPY AVAILABLE



4. FACILITATOR ASSESSMENT

(1) Present evidence to facilitator

- All evidence needs to be presented to the module facilitator
- Evidence may be written, demonstrated, discussed or pointed out in assessment reports etc

Important Note: *Your **FIRST** Problem Solving assessment will involve a discussion with the module facilitator to ensure you understand the assessment process.*
Future assessments may involve less direct contact with the facilitator if you prefer.

(2) Assessment by facilitator

The facilitator will use a ***Problem Solving Assessment*** sheet to assess all evidence.

5. DISCUSS FINAL ASSESSMENT RESULTS

COMPARE your self assessment with facilitator's assessment and discuss and negotiate any differences. This is a good opportunity to learn more about your Problem Solving processes.

6. FINAL RESULT RECORDED IN SMART

Facilitator will enter successful Key Competency result for 'Solving Problems' in SMART against the relevant module assessment activity. To get formal recognition for a Key Competency Performance Level it must be demonstrated in 2 different assessments.

* ***Important Notes for Facilitator:***

*1) Refer to **Key Competency Assessment Guide** to determine Performance Level (1, 2 or 3). Mainly determined by the following section of the **Problem Solving Assessment** form **PLUS** some other criteria.*

Execution		✓	Performance Level
Application of strategies	Applies an established procedure	☐	1
	Selects and implements a strategy from three available	☐	2
	Adapts an existing, or creates a new, procedure	☐	3

2) Finally, please add a brief explanation of the student's performance at the bottom of the entry in SMART to support this Key Competencies achievement.

BEST COPY AVAILABLE

Appendix 3: The problem-solving inventory (modified)

Name: _____

Problem-Solving Inventory

Please read each of the statements below then check the box on the right that you think most closely applies to your problem-solving in Electronics and Information Technology. There are no right or wrong answers.

	Almost Always	Often	Seldom	Almost Never
When a solution to a problem was unsuccessful, I examine why it didn't work.	☐	☐	☐	☐
When I am confronted with a complex problem, I develop a strategy to collect information so I can define exactly what the problem is.	☐	☐	☐	☐
When my first efforts to solve a problem fail, I become uneasy about my ability to handle the situation.	☐	☐	☐	☐
After I have solved a problem, I do not analyse what went right or what went wrong.	☐	☐	☐	☐
I am usually able to think up creative and effective alternatives to solve a problem.	☐	☐	☐	☐
After I have tried to solve a problem with a certain course of action, I take time and compare the actual outcome to what I thought should have happened.	☐	☐	☐	☐
When I have a problem, I think up as many possible ways to handle it as I can until I can't come up with any more ideas.	☐	☐	☐	☐
When confronted with a problem, I consistently examine my feelings to find out what is going on in the problem situation.	☐	☐	☐	☐
I have the ability to solve most problems even though initially no solution is immediately apparent.	☐	☐	☐	☐
Many problems I face are too complex for me to solve.	☐	☐	☐	☐
I make decisions and am happy with them later.	☐	☐	☐	☐
When confronted with a problem, I tend to do the first thing that I can think of to solve it.	☐	☐	☐	☐
Sometimes I do not stop and take time to deal with problems, but just kind of muddle ahead.	☐	☐	☐	☐
When deciding on an idea or a possible solution to a problem, I take time to consider the chances of each alternative being successful.	☐	☐	☐	☐
When confronted with a problem, I stop and think about it before deciding on the next step.	☐	☐	☐	☐
I generally go with the first good idea that comes into my head.	☐	☐	☐	☐

Please turn over...

BEST COPY AVAILABLE

	Almost Always	Often	Seldom	Almost Never
When making a decision, I weigh the consequences of each alternative and compare them with each other.	☐	☐	☐	☐
When I make plans to solve a problem, I am almost certain that I can make them work.	☐	☐	☐	☐
I try to predict the overall result of carrying out a particular course of action.	☐	☐	☐	☐
When I try to think up possible solutions to a problem, I do not come up with very many alternatives.	☐	☐	☐	☐
Given enough time and effort, I believe I can solve most problems that confront me.	☐	☐	☐	☐
When faced with a novel situation, I have confidence that I can handle problems that may arise.	☐	☐	☐	☐
Even though I work on a problem, sometimes I feel like I am groping or wandering, and not getting down to the real issue.	☐	☐	☐	☐
I make snap judgements and later regret them.	☐	☐	☐	☐
I trust my ability to solve new and difficult problems.	☐	☐	☐	☐
I have a systematic method for comparing alternatives and making decisions.	☐	☐	☐	☐
When confronted with a problem, I usually examine what sort of external things my environment may be contributing to the problem.	☐	☐	☐	☐
When I am confused by a problem, one of the first things I do is survey the situation and consider all relevant pieces of information.	☐	☐	☐	☐
Sometimes I get so charged up emotionally that I am unable to consider many ways of dealing with problems.	☐	☐	☐	☐
After making a decision, the outcome I expected usually matches the actual outcome.	☐	☐	☐	☐
When confronted with a problem, I am unsure of whether I can handle the situation.	☐	☐	☐	☐
When I become aware of a problem, one of the first things I do is to try to find out exactly what the problem is.	☐	☐	☐	☐

Thank you for your participation

Appendix 4: The problem-solving assessment instrument

The first page of the problem-solving assessment instrument is referred to in appendix 2 as the 'problem-solving assessment sheet' and the 'problem-solving assessment form'.

Demonstration of Problem Solving Performance

Name _____ **ID** _____

Module _____ **Task** _____

Facilitator _____ **Date** _____

(SA = Self Assessment, V = Facilitator Validation)

Defining the Problem		SA	V
Forms a correct understanding of the problem	Misunderstands the problem	☐	☐
	Forms a partial understanding using some given information	☐	☐
	Forms a complete understanding using all relevant factors	☐	☐
Recognises significance of given information or the need for new information	Fails to recognise the significance of information that is given	☐	☐
	Recognises the significance of most information that is given	☐	☐
	Recognises the significance of all information that is given	☐	☐
	Identifies specific additional information required	☐	☐
Recalls relevant information	No recall, recalls irrelevant information, inaccurate recall of information	☐	☐
	Accurately recalls relevant information as isolated elements	☐	☐
	Accurately recalls relevant information as integrated elements	☐	☐
Sets a realistic goal	Does not identify a clear goal	☐	☐
	Identifies relevance of given goals	☐	☐
	Establishes a clear goal	☐	☐

Planning an Approach		SA	V
Plans an approach to the problem	Does not engage in planning activity	☐	☐
	Undertakes enough planning to begin the solution	☐	☐
	Plans all stages required to achieve the goal	☐	☐
	Considers several alternative approaches and selects the best plan to achieve the goal	☐	☐
Recalls previous relevant or similar problem tasks	Does not refer to previous problems	☐	☐
	Recalls previous problems that may be relevant	☐	☐
Identifies appropriate sub-goals	Does not identify sub-goals	☐	☐
	Breaks the task into smaller sub-goals	☐	☐
Sets an appropriate time frame	Does not consider time frame	☐	☐
	Estimates how long the problem solution should take	☐	☐
	Integrates other commitments into time frame established for solving this problem	☐	☐

For each indicator, eg 'Sets a realistic goal', select the criterion description that most closely matches task performance and check its box. If you believe that an indicator has been performed, but there is no direct evidence of it, leave the boxes for that indicator blank.

BEST COPY AVAILABLE

Carrying out a Plan		SA	V
Begins to follow the set plan	Begins to work on the problem without a clear system	☐	☐
	Works systematically, but without reference to plan	☐	☐
	Works systematically and follows set plan	☐	☐
Activates relevant knowledge	Does not activate relevant knowledge or activates incorrect information	☐	☐
	Activates some, but not all, relevant knowledge	☐	☐
	Activates and uses accurate relevant knowledge	☐	☐
Application of strategies Key Competency Performance Level	1. Applies an established procedure	☐	☐
	2. Selects and implements a strategy from those available	☐	☐
	3. Adapts an existing, or creates a new, procedure	☐	☐

Monitoring Progress		SA	V
Checks progress towards goal	Does not check progress against sub-goals or final goal	☐	☐
	Periodically checks solution progress against sub-goals and final goal	☐	☐
Responds to unexpected problems along the way	Does not recognise unexpected problems	☐	☐
	Identifies unexpected problems	☐	☐
	Diagnoses causes of unexpected problems	☐	☐
	Makes adjustments to rectify suspected causes of unexpected problems	☐	☐
Reviews original plan	Does not review original plan	☐	☐
	Reviews original plan	☐	☐
	Reviews and revises original plan or verifies that original plan is appropriate	☐	☐
Checks original understanding and definition of problem	Does not check original understanding and definition of problem	☐	☐
	Reviews understanding and definition of problem	☐	☐
	Reviews and revises understanding and definition of problem or verifies that original definition is correct		

Reflecting on the Result		SA	V
Reviews efficiency and effectiveness of problem approach	Does not reflect on solution efficiency and effectiveness	☐	☐
	Reviews efficiency and effectiveness of solution strategy	☐	☐
	Identifies improvements to solution strategy	☐	☐
Compares current problem with previously encountered ones	No comparisons with previous related tasks	☐	☐
	Compares current task with previous ones	☐	☐
	Notes ways in which current experience might have helped past problems	☐	☐
Anticipates situations in which current problem approach might be useful	No anticipation of future applications of solution strategy	☐	☐
	Considers future broader applications of solution strategy	☐	☐

For each indicator, eg 'Sets a realistic goal', select the criterion description that most closely matches task performance and check its box. If you believe that an indicator has been performed, but there is no direct evidence of it, leave the boxes for that indicator blank.

BEST COPY AVAILABLE

Scoring Instructions

For each indicator, select the level of performance that matches the student's performance. The first option for each indicator is the 'did not perform' level. Other options indicate levels above this.

If you believe that the student has implicitly performed on this indicator, but has not presented obvious evidence, leave the boxes for that indicator blank.

Defining the Problem

Forms a correct understanding of the problem. How complete is the student's understanding of the problem? Has s/he formed a simple understanding based on one aspect of it (say resistance using Ohm's Law) or a more complex one based on several aspects (say resistance and heat dissipation and ...)?

Recognises significance of given information. Has the student noted the significance of information that is given? Does the student recognise that the information that has been given and what they already know are not enough to enable them to proceed. They may need to check technical manuals or other reference sources to identify specific additional information.

Recalls relevant information. Does the student recognise that they already know things that are important in solving the problem? Does s/he recall that information accurately? Is the information recalled as separate pieces of information or does the student show an understanding of how it all relates together in relation to the problem task?

Sets a realistic goal. Does the student identify a given goal or establish their own clear and achievable goal? This is how they will know that they have reached an appropriate solution.

Planning an Approach

Plans an approach to the problem. Does the student reveal any planning? Sometimes this will be short term planning, ie just enough to get started but not enough to get to a final solution. This is fine for novices, because it will be easier for them to set a short term goal and to achieve that before going onto the next phase of the problem. A more experienced student will set a series of goals that will guide them from the beginning of their solution attempt to the goal.

Recalls previous relevant or similar problem tasks. Good problem-solvers think back over their past experiences to find a problem, similar to the current one, that they have encountered previously. This will help them to develop a solution strategy. Can the student tell you about similar past problems? Can they tell you in what ways past problems were both similar to and different from the current one? Can they use this to help develop a solution strategy?

Identifies appropriate sub-goals. Many problems involve sub-problems or phases in their solution attempt. Can they set a final goal and then establish milestones or sub-goals along the way?

Sets an appropriate time frame. Is the student clear about how long the problem should take to solve? Is their estimate reasonable? Do they integrate this with other commitments they may have?

Carrying out a Plan

Begins to follow the set plan. Does the student begin to work systematically? If the student has a clear plan, does s/he show that they are following it? In cases where experienced students did not seem to develop a plan, is there evidence that they are working systematically?

Activates relevant knowledge. If the student needed to recall information, has s/he been able to apply that knowledge correctly in the current situation?

Application of strategies. Did the student apply an established procedure (eg experiment procedure or flowchart)? Did they select and implement an appropriate strategy from those available? Did they adapt or manipulate an existing, or create a new, problem solving process?

BEST COPY AVAILABLE

Monitoring Progress

Checks progress towards goal. Does the student pause, or do they give any indication that during the solution attempt, they paused and checked their progress against sub-goals or the final goal to see that they were getting closer to a solution?

Response to unexpected problems along the way. If an unexpected problem arose, and they were not able to proceed as expected, do they show any evidence that they have looked for the cause? For example, do they show the need to retest existing components or recheck a circuit design? Do they seek additional information to verify their understanding of the problem? If the solution is going according to plan, especially if they are following a set procedure, this indicator may not be apparent.

Reviews original plan. If the original plan does not appear to be working, does the student review it and adopt a new approach or set new goals? If the solution is going according to plan, does the student review the plan to verify that it cannot be improved.

Checks original understanding and definition of problem. If the student has found that the solution attempt is not working, do they go back to the definition of the problem and pause to think again about the nature of the problem? Here, a student who had formed an incorrect understanding or definition of the problem might think about other possible dimensions of the problem and redefine the problem.

Reflecting on the Result

Reviews efficiency and effectiveness of problem approach. Having solved the problem, has the student considered how they might have been more efficient? Would the student make any major changes to the way they went about this activity if they had to do it again?

Compares current problem with previously encountered ones. Does the student compare the just-completed problem with others that had been attempted earlier? Can they identify techniques in this solution that would have helped with earlier problems? Were there techniques that were used previously that would have helped in the current problem?

Anticipates situations in which current approach might be useful. Does the student show evidence that what they have just done could be used in other real world situations? Can they identify a specific task where the current approach is likely to be applicable?

BEST COPY AVAILABLE

Appendix 5: Student evaluation comments

What follows are comments made by students during the evaluation that they were asked to complete.

Student 1: K C not only make future job employers aware of your skills but it makes you aware of your skills and gives you the extra confidence in a job interview situation.

Student 2: I did find my first attempt the most difficult and time consuming but now there is greater information available and I have completed a few more I find them an easy way to improve my 'soft' skills and analyse my methodologies.

Student 3: When they were first introduced I did a few but then completely forgot about them due to the complexity of understanding the performance levels. Now due to the fact that [facilitator] is insisting we gain some while doing [Course name], I have gained some more. It is only with [facilitator's] help and prompting that I have gained a little understanding of the terminology that decides the level to go for. When it gets too difficult to understand what the different levels are about, I just decide not to go for them. Doing these assessments have partly helped me understand and improve my skills because it put a name to them but I did not improve the skills, in the ones I have done anyway. I find it difficult to differentiate between some of the levels.

> **Facilitator's response:** The issue of differentiating between performance levels was discussed further with this student and was suggested that the new Problem Solving Assessment sheet is designed to be easier to follow and understand. It has less of a focus on performance levels. The student said he would like to try this new assessment sheet in the near future. I said it would be very helpful feedback to hear what he thinks about it in light of his current comments. He indicated that recent facilitator support focussed directly on key competencies has been of enormous help. Through discussion he also acknowledged that the new assessment sheet may be more useful in helping him understand and improve his skills.

Student 4: I got kicked in the guts with KC a few times. I think that its too hard and pointless to do. I think Experience will win over KC. I don't think I will be doing anymore, because it takes too much time.

> **Facilitator's response:** This particular student participated in one Problem-solving Assessment via an informal interview. Being the first time it is always somewhat lengthy (approx. 1 hour) in order to ensure the student develops a clear understanding of the whole assessment strategy and the potential benefits etc etc. This particular student unfortunately was not successful in his Problem-solving Assessment on this occasion as it became evident that he had made some fundamental errors. He misunderstood the actual problem and his execution and monitoring were flawed by the fact that his strategies did not in fact address the real problem. The outcome was that this assessment exercise was identified as a learning experience which should better prepare the student for future assessments. He was encouraged to ask questions, seek assistance and consider future assessments—to which he replied that he would. In fact he actually named the one he intended to do next. I believe this is an important opportunity for this student [to]

acknowledge where he is currently at in terms of 'problem solving' and also 'self assessment' (as identified through this assessment discussion) in order to improve and develop these important skills. I have also just learned that this student has attempted about 3 other KCs. One was successful but the others were not. The facilitator involved indicated that this was due to a lack of ability to provide suitable evidence. That facilitator has likewise seen this as an opportunity to assist this student to develop KC and self-assessment skills.

Student 5: I think it is great that this Tafe campus offers these great recognisation of skills.

Student 6: Understand the importance of key competencies. At the moment I'm a new student trying to work my way through the first prerequisite modules.

Student 7: The only Key Competency Assessment I did did leave me feeling more confused than informed. However, since all the guest speakers etc I have a better idea of what it's all about.

In response to the following question: Is the assessment process clear and easy to follow? ...

Yes, after the 1st assessment with [facilitator] [to induct students into the process].

First Problem-solving Assessment should be made compulsory for the Diploma. The second assessment can be kept for the Cert [?].

Student 8: Does require a bit of work and discussion to follow through with the principle of the key competencies. e.g. Having it explained by a facilitator.

Student 9: As soon as I have time I will get around to it.

Student 10: Maybe more info can be given on how long it takes to get KC's. I wasn't ever told that you have to apply before starting a module. I think that they should be done at the easy stages of the courses.

> **Facilitator's reply to S10:** Thanks for taking the time to provide some feedback on key competencies.
>
> It is truly essential for us to get this feedback in order to improve the system for all students.
>
> Just to clarify something for you ... You actually DON'T have to apply for Key Comps before your start the module—some facilitators may encourage you to think about it and advise them of your intentions at the beginning of the module BUT it is definitely not a requirement.
>
> I'm sorry if you were given that impression or if that is what was told to you.
>
> It is fine to decide immediately after you complete the module assessment activity to apply for key competencies. However, it obviously helps if you plan ahead and can be mindful of Key Competency processes as you work through the module assessment activity.
>
> I am happy to discuss this with you further if you like.
>
> Thanks again!

Student 11: I am self employed and I can not see myself work for someone else but any key competencies will help in studies and my self confidence.

Student 12: Having the same process as the new Problem-solving Assessment available for the other key competencies would be beneficial.

Student 13: Some of the wording in the key competency checklists are difficult to understand and it is hard to know what exactly they require.

Student 14: The only issue I could comment on is that at the time I completed the assessment it was unclear exactly what parts of the curriculum were applicable to key competency assessment.

Student 15: It was difficult to know when a key competency could be attempted.

Student 16: Some of the wording was too technical, I think that the use of plain words would make the process of choosing which box easier to decide.

Student 17: Very helpful in breaking down the problem-solving method into identifiable sections. Helped me better understand the process.

Student 18: The original way of assessment for key competencies is not that clear on what is exactly required for each level of competency. The new 'problem-solving' trial has helped much more in that it breaks down each section of the process and allows for detailed discussion each time.

Student 19: I think the key competencies is a good way of being recognised for things that you do but are not necessarily recognised in any other way.

- compare to next assessment might be helpful
- be aware of areas for improvement
- How can we improve areas?

Student 20: I reckon key competencies assessment is a good idea because it allows students to get recognised for the work they do.

The National Centre for Vocational Education Research is Australia's primary research and development organisation in the field of vocational education and training.

NCVER undertakes and manages research programs and monitors the performance of Australia's training system.

NCVER provides a range of information aimed at improving the quality of training at all levels.

ISBN 1 74096 167 6 web edition

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

X	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").